

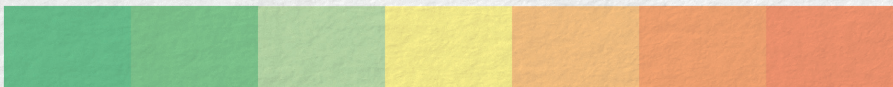
DATA

DATA

AND EVEN MORE

DATA:

Empowering users to make
well-informed decisions about
online privacy



Susanne Barth

**DATA, DATA, AND EVEN MORE DATA:
Empowering users to make well-informed
decisions about online privacy**

Susanne Barth

Susanne Barth

Data, data, and even more data: Empowering users to make well-informed decisions about online privacy

ISBN: 978-90-365- 5141-0

DOI: 10.3990/1.9789036551410

Cover design by Yasmin Katlich, persoonlijkproefschrift.nl

Layout and design by Yasmin Katlich, persoonlijkproefschrift.nl

Printed by Ipskamp Printing | proefschriften.net

Copyright © 2021 Susanne Barth, The Netherlands | Germany

All rights reserved. No parts of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author.

DATA, DATA, AND EVEN MORE DATA:

Empowering users to make well-informed decisions
about online privacy

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. ir. A. Veldkamp,
on account of the decision of the Doctorate Board,
to be publicly defended
on Friday 09 April 2021 at 12.45 hours

by

Susanne Barth

Born on 20 January 1982
in Herdecke, Germany

This dissertation has been approved by:

Supervisors

Prof. dr. M.D.T. de Jong

Prof. dr. M. Junger

Graduation Committee

Chairman/secretary Prof. dr. T.A.J. Toonen

Supervisors Prof. dr. M.D.T. de Jong

Prof. dr. M. Junger

Committee members: Prof. dr. S. Ben Allouch, University of Amsterdam/Amsterdam
University of Applied Sciences

Prof. dr. P.H. Hartel, University of Twente

Prof. dr. ing. A.J.A.M. van Deursen, University of Twente

Prof. dr. B. van Lier, Steinbeis University Berlin/University of
Applied Science Rotterdam

Prof. dr. phil. M. Ziefle, RWTH Aachen University

Contents

Preface		9
Chapter 1	Introduction	11
Chapter 2	The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review	37
Chapter 3	Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources	71
Chapter 4	Lost in privacy? Online privacy from a cybersecurity expert perspective	97
Chapter 5	Toward an understanding of online privacy perceptions: Using the Q-sort method to identify different user perspectives	119
Chapter 6	Understanding online privacy - A systematic review of privacy visualizations and Privacy by Design guidelines	141
Chapter 7	Privacy Rating: A user-centered approach for visualizing data handling practices of online services	189
Chapter 8	General discussion	219
References		237
Appendices		279
Summary	(in Dutch)	295
Biography		303
Acknowledgements		309

Preface

Why anyone should care about privacy, privacy engineering or data at all

It's time to serve humanity.

Humanity is people.

Humanity is empowered stewardship of our surroundings—

Our universe, planet, and future.

Humanity is described by data;

Data about humans;

Data about all things human.

Data is not humanity;

Data tells a story;

Data is leverage;

Data is not power.

Humanity can capture data.

Data cannot capture humanity.

It's time to serve humanity.

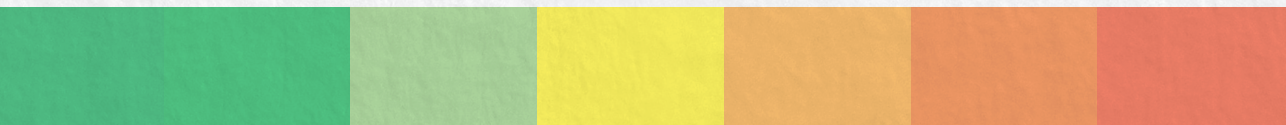
There is no one else.

We are already past due.

This is the paradox in which the privacy researcher [original: engineer] discovers, inspires, and innovates.

Let's begin.¹

¹ Boanabeau (2014). Foreword to Finneran Dennedy et al. (2014), *The Privacy Engineer's Manifesto. Getting from Policy to Code to QA to Value*. Apress.



1

Introduction

“Privacy is not only an arbitrary cultural and legal concept. It is an anthropological constant and a psychological necessity. It is a complex process of selectively managing access to one’s self. Without a minimum of privacy people can’t survive” (Boehme-Neßler, 2016, p. 222).

In most Western countries, privacy is considered a human right that serves humanity and the ‘stewardship of our surroundings’ —including privacy with respect to our personal data (Solove, 2008). This aspect of personal data refers to the informational dimension of privacy and the control individuals have over information pertaining to their data (Roessler, 2005). In the remainder of this thesis, online privacy (often referred to as ‘privacy’) is understood as a subtype of informational privacy with personal data “relating to an identified or identifiable natural person” (GDPR, Art. 4) and subjected to privacy threats and protection measures (The European Parliament and the Council of European Union, 2016). Furthermore, personal data, special categories of personal data and the processing of personal data are understood as defined in the GDPR Art. 4 and 9.

The *psychological* need for information privacy is deeply rooted in human consciousness (Holvast, 2009). It evolves in the early years of childhood and is *culturally* independent, at least to a certain degree (Schütz & Friedewald, 2011). At a very young age, children begin shielding themselves and their intimate areas; they want to play alone without anyone keeping an eye on them, and they understand the concept of secrets. During this phase, children experience privacy by keeping things to themselves until they decide to share and communicate any such information with others. Such initiatives are the first steps toward developing an autonomous ‘self’ concept and exercising control (Introna, 1997; Kupfer, 1987).

From a *legal* perspective, Warren and Brandeis (1890) defined privacy as ‘the right to be left alone’, which essentially means the protection of thoughts, sentiments and information originating from inter- and intrapersonal communication. Although Warren and Brandeis defined privacy in an offline context, the definition can arguably also be applied to online privacy needs—although defining and maintaining the boundaries in this context is much more difficult. Almost a century later, Westin defined privacy as “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, 1967, p. 7). Albeit many other definitions of information privacy exist, for instance, privacy as a bilateral contract between parties within a given community and context (K. Martin, 2016a) or shaped by *social* boundaries and context-specific norms (Nissenbaum, 2004). Because of its context dependency, formulating a one-size-fits-all definition of privacy is almost impossible—and for all intents and purposes, should be understood in terms of common ‘core characteristics’, norms and values (Nissenbaum, 2004; Solove, 2002, 2008).

Essentially, the umbrella term privacy centers around the concepts of freedom and self-determination or in more *technical* words—access and control—in both

online and offline contexts. From an *ethical* perspective, the critical point lies to a lesser extent in access to information but more in self-determined control over access, governed by social norms and contextual factors. In some situations, disclosure of information can be beneficial and acceptable, whereas in other situations, it can be damaging (Acquisti et al., 2016; Holvast, 2009; Nissenbaum, 2004; Roessler, 2005; Sax, 2018). Some researchers considered the seemingly unlimited possibilities for data gathering and analysis to be the most pervasive in human history (Introna, 1997), while other scholars even called the technological developments of recent times “the uncontrolled panopticism” (E. J. Smith & Kollars, 2015; p. 160). It should be noted that at the time, the increasing scope of real data analytics made possible through things such as the ‘Internet of Things’ and advances in artificial intelligence were not yet factored in. Although such technologies have their merits and are responsible for an unprecedented rise in quality of life, they also have their downsides. On an individual level, unlimited self-disclosure and data gathering can, for instance, lead to undesirable identification, reputational damage, financial loss, dissemination of sensitive information, loss of autonomy and psychological illness (Michelfelder, 2001; Schütz & Friedewald, 2011).

On a *societal* level, technologies can, for instance, be misused for the corruption of elections, criminal infiltration, or undermining the credibility of democratic institutions (Manheim & Kaplan, 2019). Therefore, the right to control data flows online and the right to be forgotten are central to the discussion around online privacy and data protection. To safeguard autonomy, humans must be able to restrict access to certain information about themselves (A. L. Allen, 2012), and the absence of data protection threatens autonomy (Boehme-Neßler, 2016; Schwartz, 1999). Privacy management is a continuous tension between disclosure and withholding information, mediated by technology and context. To understand privacy concerns and the management of privacy boundaries, it is important to consider not only the individual level but also the “whole of the social and institutional setting in which technologies are deployed” (Palen & Dourish, 2003, p. 135).

However, the questions around privacy did not change with new technologies, but privacy itself became more complex and difficult to control and implement. Some researchers even argue that protecting privacy will be one of “the greatest civil liberty issues of the new millennium” (Akdeniz, 2002, p. 194; Boehme-Neßler, 2016). The abovementioned discussion shows that information privacy is not only a fluent concept in itself but also approached differently depending on the perspective from which it is being observed. This fluency and the various

perspectives on the topic make information privacy a ‘concept that is still in disarray’ (Solove, 2008) and difficult to define (Pavlou, 2011).

To understand information privacy as a multidisciplinary concept, relevant perspectives on information privacy from different disciplines will be presented in the remainder of this chapter. First, a historical view on information privacy will be presented to show how the perception of information privacy developed over time, paying particular attention to the ramifications of the technological innovations that now play a part in everyday life. Second, not only has the perception of information changed over the course of history, the legal definitions and the measures implemented to protect information privacy have also undergone changes that will be discussed thereafter. Third, this will be followed by presenting a technical perspective on the topic that shows how online privacy is instrumentally approached. To discuss the relationship between information privacy and the user itself, two further perspectives are chosen. To understand the power relationship between users and data holders in a data-driven market, information privacy will be approached from an economical perspective. Furthermore, in the discussion of online privacy from a social perspective, the formation of users’ privacy perception, the processes that guide users’ online behavior and the role of the context that determines users’ privacy preferences are scrutinized.

1.1 The right to be left alone: A historical perspective on information privacy

*“We must conclude that we are increasingly going to live in a surveillance society in which almost everything about our lives will be known”
(Holvast, 2009, p. 14).*

Long before the birth of mass media, privacy was closely associated with the right of property. This essentially meant that the protection of the home as the castle and the protected area should not be compromised by trespass (Semayne’s case), albeit often depending on the social and economic status of citizens. It still took a while before privacy was considered a right to protect the individualism and autonomy of all human beings, at least in Western societies (Hafetz, 2002; Vickery, 2008; Vincent, 2016).

Although the invention of affordable printed media in the 17th century is considered the birth of mass media, the discussion around privacy gathered steam with the launch of photography, motion pictures and telephony in the

late 19th century. Warren and Brandeis (1890) critically observed those technical developments and postulated in their landmark book ‘the right to privacy’ that modern life and civilization advances increase publicity susceptibility. Through mass media, the distribution of personal information exceeds the private sphere of family and friends to reach the broader public, often without having given explicit consent to this level of distribution. Hitherto, the focus of privacy was more on the balance between privacy and free speech, the reasonability of public interest and the exposure of private individual information to a wider public through mass media (Gavison, 1992). This is not to say that these foci are no longer relevant, but the development of the administrative state (e.g., census and income tax), economic growth and organizational innovations (e.g., mass production, new marketing and analytic possibilities) in the early 20th century introduced a new dimension to privacy: the collection and analysis of an increasing amount of information about people (Barbas, 2012).

Although those developments eventually changed the perception of public opinion about the invasion of privacy as the first citizens raised their voices, these times were still characterized by high trust in the government and economic sector (Barbas, 2012; H. J. Smith et al., 2011). However, in the second half of the 20th century, large-scale civil protest and dissatisfaction about data gathering practices, governmental surveillance and eavesdropping questioned the boundaries of individual privacy. More specifically, in the 1970s, with the introduction of personal computers and commercial mobile phones, the potential threats of new technologies were recognized (Barbas, 2012; Brenton, 1964; H. J. Smith et al., 2011). At the end of the 20th century, the discussion around information privacy further intensified with the establishment of networked systems and sophisticated database capabilities, e.g., direct marketing (H. J. Smith et al., 2011).

A new wave of privacy concerns was raised in the 2000s with the mainstream use of the internet, the ubiquitous adoption of cell phones and text messaging. These new technologies opened up new possibilities for monitoring, data mining, profiling and sophisticated e-commerce initiatives (Boritz & No, 2011; H. J. Smith et al., 2011). This was followed by cloud computing and the emergence of social media platforms around a decade later, the evolution from cell phones to smartphones, the rise of newer technologies such as the Internet of Things (IoT) that experienced mass adoption in approximately 2013 (although the term was already invented in the ‘90s; see, e.g., Kramp et al., 2013; Van Kranenburg, 2011), artificial intelligence (AI) advancements, social robots, or quantum computers, that all have one thing in common: vast amounts of collected data. While gathering information about others is historically not a

new phenomenon, digital technologies have created a readily accessible data-rich environment—the likes of which society has never seen. IoT applications associated with smart homes, smart cities, or e-health now link appliances that were previously autonomous, saving, sharing and utilizing the user-based data generated. This creates a ‘mega market’ with the ultimate consequences of interwoven devices connected to the internet still unknown (Kramp et al., 2013). Moreover, IoT’s AI engine aims to structure real-world data from people and to filter noise from big data to predict and deduct meaning and to draw inferences with precision transcending human power and capabilities (Manheim & Kaplan, 2019). Ultimately, however, these technologies can only find their legitimacy and success in a market that demands them. In other words, the push-pull combination of technology that *pushes* and humans that *pull* for this ever-increasing connectivity is what makes the IoT “very strong, unstoppable, fast and extremely disruptive” (Kramp et al., 2013, p. 2).

As with many aspects in human life, there are two sides to every coin, and the digital revolution is no exception: while the advantages of technological achievements initially enjoyed the limelight, the rise of the IoT, big data and sophisticated data analytics have forced the dark side of such technologies to emerge from behind the curtains. An increasing number of people have expressed concern about things such as (but not limited to) profiling, voter manipulation, algorithm-driven decision-making and persuasive computing (Manheim & Kaplan, 2019). The main difference between the data gathering today and that of the 20th century, for example, is not only the scope these technologies have made possible but also, more importantly, the current ability to store, analyze and aggregate the data gathered. This paves the way for far more intrusive personal privacy violations, and its consequences are unforeseeable, endangering the privacy of individuals more than never before (Solove, 2002). Interestingly, more than 100 years after the Warren and Brandeis era and even after the ‘cybernetic revolution’ of the 1970s (Miller, 1969), the invention of the internet in the 1990s and today’s multimedia world, lawyers, scientists and practitioners have still failed to reach an agreement on a blueprint for the principles of data protection (Anwar et al., 2018; Martín et al., 2014; Nissenbaum, 2015). One thing, however, is sure: What we can learn from former generations is to take invasion of privacy seriously (Barbas, 2012; Gavison 1992).

1.2 Challenges for data protection: A legal perspective on information privacy

“Wherever we go, whatever we do, we could easily leave behind a trail of data that is recorded and gathered together. These new technologies, coupled with the increasing use of personal information by business and government, pose new challenges for the protection of privacy”
(Solove & Schwartz, 2021, p. 2).

Technological developments, especially the emergence of mobile computing, IoT, and AI, present data protection laws with new challenges. Big data and new analysis techniques pave the way for calculating probabilities and correlations derived from vast data points, providing insights into social, economic, political or epidemiological trends, processes and problems (Boehme-Neßler, 2016; Hill, 2014).

Some scholars argue that the principle of big data—the gathering of vast amounts of information, having either a direct or an indirect relationship with personal information—is contradictory to existing laws and regulations pertaining to the data protection of users (Boehme-Neßler, 2016) or that existing law lags behind those developments (Solove, 2002, 2006). As Boehme-Neßler (2016) points out, “data protection is a further specification of human dignity that takes into consideration the special challenges and risks of the digital knowledge society” (p. 224). Data protection law must counteract the imbalance between data subjects (those who provide data) and data controllers (those receiving, collecting, storing and disseminating data). Balancing ‘informational power’ eventually benefits self-determination and controls personal data (Rouvroy & Poullet, 2009). Other researchers go a step even further and state that the possibilities of emerging technologies such as monitoring and profiling not only threaten the autonomy and control users have over their personal data but, more importantly, the autonomy of individual decision-making. Therefore, constitutional rights to information privacy and its protection need to consider the moral layer of information privacy. That is, it recognizes the relationship between autonomy, decision-making and information privacy (Michelfelder, 2001).

It is reasonable to claim that technological developments and the current state of data protection laws are not in sync, as the former is generally ahead of the latter (Manheim & Kaplan, 2019; Schwartz, 1999). The call by Warren and Brandeis (1890) for the right to privacy and the right to be left alone is intertwined with the birth of privacy legislation in the United States. Warren and Brandeis not only called for a principle for privacy protection, what they claimed

to be the ‘right to privacy’ but also demanded legislative bodies to explicitly describe the nature and extent of privacy protection measures. They denounced that common law upheld adherence to the right of property and privacy inside the confines of a building but left the “back door” open for the “prurient curiosity” of unwanted listeners (Warren & Brandeis, 1890; p. 220). They questioned the discourse about privacy that centered around a clear distinction between the private and the public sphere based on whether “the individual or the society at large were involved” (Glancy, 1979; p. 19) and emphasized the right to privacy as important for mental wellbeing. Essentially, they opened the discussion on what is currently called *control*: the right of each individual to be able to decide for themselves what is private and what can be shared publicly—and that this right is protected by law (Glancy, 1979; Warren & Brandeis, 1890).

However, the interpretation and conceptualization of such a right was—and still remains—vague (Barbas, 2012; Solove, 2002). The difficulty with information privacy law lies in the nature of privacy itself. The perception of the intimate space allocated each human a century ago has changed, as has the nature and status of a family unit or the common perceptions of private and working spaces. While technologies also develop over time, they do so historically much faster than events shaped by society. Consequently, the law encounters difficulties in closing the gaps between current data protection regulations and technological developments. Developments in AI are a striking example of this discrepancy, as the first congressional hearings on advances in this field were not held until more than a half-century after AI was recognized in science and the military (Haenlein & Kaplan, 2019; Manheim & Kaplan, 2019).

Many judicial systems, especially in the western parts of the world, have been trying to catch and maintain pace with these technological developments (Solove, 2002). The U.S. regulates privacy using sector-specific regulations, whereas the European Union tries to address privacy holistically. It has been suggested that the amendments to the U.S. Constitution pertaining to privacy are rather narrowly defined and refer to specific concepts and social constructs such as family, marriage, motherhood or child rearing and are considered on a case-by-case basis (C. Allen, 2019; Kerr, 2004; Sharp, 2013). Overall, the current patchwork U.S. law landscape is widely considered to be less than ideal in regard to understanding and applying privacy protection (C. Allen, 2019; Kerr, 2004; Sharp, 2013). It is quite possible that European data protection laws will play a significant role in future privacy dialogues within the U.S. (Solove & Schwartz, 2011), as the California Consumer Privacy Act (CCPA) recently demonstrated. Resembling the European General Data Protection Regulation (GDPR), the CCPA that came into force in 2020 is a game changer in privacy legislation, as

it provides the most essential data privacy protections in U.S. history—already inspiring other states to introduce consumer privacy legislation (C. Allen, 2019; National Conference on State Legislatures, 2020; Rothstein & Tovino, 2019). As with the European regulation on data protection, national coverage and uniformity are considered the ideal situation for information privacy legislation in the U.S. as well. However, while few experts see comprehensive federal privacy legislation on the horizon, many consider the introduction of CCPA to be a first step toward national data protection (Rothstein & Tovino, 2019).

At the European level, with the launch of the GDPR, the perspective on privacy shifted from an ‘annoying nice to have liberal value that may or may not be considered at a later stage’ (Introna, 1997, p. 274) to implementing privacy in the early stages of the design process and by default (GDPR, Art. 25). The GDPR that came into force in 2018 mandates six high-level principles for the processing of personal data. First, data processing must happen in a fair and transparent manner. This principle is closely related to the rights a user has: individuals have the right to know when personal data are collected, to gain access to their personal data, to limit data collection or to opt out completely and to request data to be amended or erased (GDPR, Art. 12-23). Second, data must only be collected for specified and legitimate purposes. This means that excessive data collection and requesting irrelevant permissions is prohibited by law. Third, data collection must be relevant and limited to the defined purpose(s) to prohibit the misuse of collected personal data. Fourth, at any time, collected personal data must be accurate, and if not, such data must be corrected or erased. Fifth, data must be anonymized whenever possible, and retention must be limited according to predefined purposes. Sixth, processed personal data must be protected by appropriate security measures (GDPR, Art. 5(1)). If the abovementioned requirements are not met, the data processing is not lawful. Furthermore, the party that gathers data is accountable for compliance with the six principles. The fines for violating the law are serious and can reach up to 4% of the total worldwide turnover in the last twelve months (GDPR, Art. 5(2), Art. 83(5)).

In short, if personal data are collected, the user must be informed about every detail pertaining to the data exchange. Data collection must be based on a legal *contract* and/or explicit *consent* (e.g., explained in privacy notices and based on privacy policies) and in accordance with a *legitimate interest*. Ultimately, data handling practices must be lawful at all times.

However, there is still a long way to go as users have yet to gain full control over their data when using current information technologies. Instead of understanding privacy at a general level, conceptualizing privacy within a

particular context and focusing on actual use of data and intended outcomes of data gathering might prove successful. To date, the balance between granularity and generalization of contexts remains unresolved (Solove, 2002). However, regulations specifying concrete conditions for processing personal data “are the decisive test for discerning whether [or not] society is aware of this price and willing to pay it. If the signs of experience are correct, this payment can be delayed no further” (Simitis, 1987; p. 746).

1.3 Beyond access control and authentication: A technical perspective on information privacy

“Computers constantly produce data. It’s their input and output, but it’s also a by-product of everything they do...they sense and record more than you’re aware of” (Schneier, 2015, p. 15).

The era of big data and unprecedented interconnectivity has allowed companies to amass datasets containing incredible amounts of personal information. By recording, storing, tracking and linking events over time, computers are able to identify patterns and gain insights into the individual preferences, interests or personal characteristics of users. As a result, behavioral tracking—and the profiling it makes possible—stand to become one of the most serious threats to privacy ever known.

Techniques such as browser cookies and tracking ‘pixels’ allow monitoring the behavior of individual users across the internet. Cookies were designed to allow websites to recognize visitors to provide features such as automatic login, shopping carts, language settings, or saving preferences. However, advertising companies are increasingly using this technology on a large scale to record the activity of users across a large number of websites (Castelluccia, 2012; Nikiforakis et al., 2013). They do so by asking website owners to insert a small script (also known as a ‘tracking pixel’) into their pages in exchange for usage statistics. This allows advertising companies to create surprisingly accurate profiles of individuals based on their online behavior, potentially including information they might not even be aware of themselves (Gutwirth & Hildebrandt, 2008). These profiles are usually used to target ads, but they are also sold and shared, enabling discrimination and manipulation (Lyon, 2003; Zarsky, 2019). This kind of tracking is often invisible to users. Even if cookies are disabled or blocked, ‘fingerprinting’ allows tracking scripts to identify individuals based on specific configurations that are unique to the browser. This allows tracking users between

websites without their consent and little chance for opting out (e.g., Nikiforakis et al., 2013). Even if users gave consent to data gathering by accepting cookies, implicit techniques such as data mining can produce unforeseen data sets without authorization of the user, e.g., inferences about preferences or search patterns (Rezgui et al., 2003).

Ubiquitous advertising goes a step further than solely personalizing advertisements based on online profiles (Pennekamp et al., 2017; Spensky et al., 2016). As the extension of the user's body, smartphones enable companies to draw conclusions on users' locations, vital signs, thinking patterns and even emotional inclinations to develop even more complete profiles (Castelluccia, 2012; Shklovski et al., 2014). In mobile computing, permissions are needed for the functionality of an app; and—in theory—permissions are also meant to protect the privacy of a mobile app user. Although current permission systems have undergone several design iterations to make them more understandable and transparent, they remain opaque, and both users and developers have trouble understanding their meaning (B. Shen et al., 2021). Consequently, users often click through permission requests without paying much attention to the entailed consequences, and developers ask for more permissions as they are reasonably entitled (Benton et al., 2013; Felt et al., 2012; Zhauniarovich & Gadyatskaya, 2016). Furthermore, the business model behind mobile computing encourages developers to overprivilege apps to generate more user data for tailored advertisements and/or profiling (Khatoun & Corcoran, 2017; V. F. Taylor & Martinovic, 2016).

Exposure to privacy risks pertains not only to smartphone apps. In contrast, it is a serious problem for many online applications and services, as users' current location, e-mail address and name are the details most often shared with third-party domains (Zang et al., 2015). For instance, dating apps seem to store private messages and track the location of their users. Some of these even leak private images that might compromise users' privacy (Farnden et al., 2015). Participatory sensing apps process sensitive user data such as location or time information, heart rate or blood pressure. Geolocation information, linked to pictures taken with the smartphone or obtained from navigation apps, allows the localization of users, the revealing of moving patterns and the drawing of inferences from static locations that are visited frequently (e.g., home addresses). Reality mining (e.g., offering advertisements in real time based on the identified location, drawing inferences on health status based on moving patterns), mobile online social networks (e.g., Foursquare; publicly available information) and ubiquitous voice control (e.g., detection of voice patterns) bring a new dimension

to behavioral tracking and allow the building of even more comprehensive user profiles (Castelluccia, 2012; Friedland & Sommer, 2010).

Cybersecurity is a critical element in the protection of personal data. In fact, privacy relies on the ability of software and systems to keep sensitive data secure from malicious or unauthorized access (access control and authentication). While privacy has to do with the collection, usage, and sharing of information, cybersecurity is commonly understood in terms of confidentiality, integrity and availability (CIA principles). Confidentiality is about restricting access to authorized individuals. Integrity has to do with preventing unauthorized modification to the data. Availability is about ensuring timely and reliable access to data. Weak security results in leaks and data breaches that expose large amounts of personal information. To address this and to be able to share useful data, sensitive information is often anonymized by removing personal identifiers or by replacing them with pseudonyms. However, this does not always prevent the reidentification of individuals (Sweeney, 2002). Many data owners have limited knowledge about proper anonymization and assume that anonymity is equivalent to solely removing explicit personal identifiable information (PII, e.g., name, home address, ID or telephone number; Sweeney, 2002). However, even without those explicit identifiers, data can be reidentified because of unique combinations of datapoints that can be traced back to an individual user. For instance, based on ZIP code, gender and date of birth, 87% of the U.S. population could be uniquely identified, even without knowing their name or further information (Sweeney, 2002). Furthermore, linking two datasets to each other can also reidentify anonymized datasets (linking attack). Sweeney (2002) showed that having a dataset containing medical records from patients and a voter registration list from citizens of the same city allows reidentification of individuals by matching shared attributes.

Statistical databases often contain demographical information about individuals. Reidentification of individuals and its misuse put their privacy at risk (Rezgui et al., 2003). Sometimes, releasing datasets is legally mandatory (e.g., health records in the U.S. or national census), beneficial for society (e.g., research on epidemics such as the COVID-19 crisis), or simply happens because of economic reasons (e.g., advertising market). However, for whatever reason datasets are released, this should not happen without protecting the identity of data subjects. Anonymization is considered one of the most important privacy protection measures. A method to protect the privacy of data subjects in datasets is implementing the k -anonymity protection model (Sweeney, 2002). Suppression is one possibility to guarantee k -anonymity, which means that sensitive information is not released and replaced by a place holder. However,

suppression affects the quality of a dataset negatively and can even render it useless. Therefore, looking for quasi-identifiers in a dataset (e.g., date of birth, place of birth, profession, home address or age) and generalizing those identifiers is a better solution (e.g., age classification in ranges, replacing the date of birth with a district or the exact job title with a sector; Samarati & Sweeney, 1998).

At the system level, privacy-enhancing technologies (PETs) are designed to provide anonymity and to protect against behavioral profiling. Technology-enabled solutions for privacy preservation can be divided into three groups: 1) client-server-based solutions (automatic negotiation, e.g., P3P and encryption), 2) server-based solutions (firewalls, VPN network connection) and 3) client-based solutions (e.g., personal firewalls, trace removers or anonymizers, e.g., onion routing). In addition to technical mechanisms, privacy protection can also be regulated by law (e.g., GDPR and CCPA) or self-regulated (e.g., certified privacy policies). Privacy by Design (implementing privacy as a fundamental system requirement) and privacy certifications (e.g., TRUSTe label) shift the burden of privacy protection from users toward developers (Rezgui et al., 2003). Tools such as MockDroid (Beresford et al., 2011) or AppFence (Hornyack et al., 2011) are designed to protect users' privacy while using mobile apps. These privacy protection tools send false data in response to app requests but do not protect against third-party API calls. Furthermore, similar to do-not-track settings on websites and providing an opt-out function on mobile apps might safeguard mobile users' privacy. While Google and Apple have already implemented initiatives to include prevention features ('Limit Ad Tracking'), those settings do not protect against apps operating outside the operating system network (Zang et al., 2015).

Although PETs are considered powerful technology-based tools for protecting users against privacy-invasive practices, add-ons cannot address the root of the privacy problem without mandatory regulations. Therefore, the "symbiotic relationship between privacy and security" (Rachovitsa, 2016, p. 399) must be seen as a core design requirement, aligning with privacy by design guidelines such as minimizing data collection and control over personal data (Castelluccia, 2012). Providing transparency and informing the user (notice and consent) could hold the key to lowering the effectiveness of privacy infringement techniques. Despite 'hiding' information about sharing practices in privacy policies, notifying users about data handling practices a priori before using an online service is a first step in the direction of 'deliberately architected privacy' (Schneier, 2009).

1.4 What data are worth: An economical perspective on information privacy

“Solving the privacy problem means to find a balance between information sharing and information hiding that is in the best interests of data subjects but also of society as a whole” (Acquisti, 2010, p. 42).

The emergence of big data marked the dawn of a new economic era, characterized by the large-scale collection and monetization of information. A data economy without data cannot function as user data is the currency of such an economy. Thus, data must be gathered to keep the data economy running. However, this must not be based on a one-sided relationship where data are obtained without restrictions and little transparency (Nissenbaum, 2011). Arguably, keeping the relationship between the data subject and data holder in balance is one of the major challenges of today’s world. New technologies enable consumers to cross the border between offline and online life, and the economic landscape changes from consumers *consuming* information to consumers *producing* information. This allows the gathering and aggregating of consumer information in numerous settings, in incredible amounts, with wide coverage and substantial economic value. Consequently, as the new currency, personal data characterize the economic *power* between the data subject and data recipient. Three main aspects seem to be particularly important when understanding and assessing the economics of privacy and the value of data, namely, (1) the context dependency of privacy, (2) the weighing of benefits for society versus individual risks, and (3) the power relationship between user and data holder (Acquisti et al., 2016).

First, privacy issues are context-dependent, and therefore, a unified economic theory of information privacy is difficult to define. According to the theory of contextual integrity (Nissenbaum, 2004, 2011), users’ perception of online privacy and preferences for disclosing or withholding information differ from one context to another. What a user is willing to disclose in one particular situation may not apply to other situations. The disclosure of information is always governed by certain norms of information flow. Those norms can be explicitly defined or implicitly determined and relate to the principles of *appropriateness* (what information is appropriate to disclose in a particular situation) and *information distribution* (what information is actually distributed and is it in compliance with contextual norms). Contextual integrity—and therefore privacy—is breached if either one of the principles or both are violated (Nissenbaum, 2004, 2019).

Second, whether privacy protection initiatives are beneficial for the individual and society or if they limit individual and societal welfare also depends on the

context. Recent studies on the COVID-19 epidemic show that data mining provides insights that are beneficial for managing this health crisis worldwide. For instance, machine learning algorithm techniques classified COVID-19 genomes, predicted survival probabilities and detected potential drugs (Ayyoubzadeh et al., 2020). In this global health crisis, large-scale patient data can be beneficial for society, as they help to better understand outbreaks, enhance diagnosis accuracy and improve therapeutic effectiveness (Alimadadi, et al., 2020). However, while profiling and data mining can have a significant value on a societal level, such techniques must be used in a responsible manner, and the data gathered must be kept to the minimum possible to avoid risks for the individual, e.g., discrimination, deindividualization and information asymmetries (Schermer, 2011). This example demonstrates that personal data clearly have value. However, the benefits to society and the risks to individuals must be balanced with the economic power among involved parties. If information is *power*, the balance is reflected in the *control* one has over personal information (Acquisti et al., 2016).

Third, market interactions pertaining to personal data seldom take place with the provision of fully informed consent—a situation that fosters asymmetry in the power relationships between the user and data holder. Biased decision-making takes place as consumers encounter difficulties assessing the consequences of data sharing, are ignorant of the threats to their personal data, lack the knowledge on how to protect personal data and are guided by nonrational cognitive processes (Acquisti, 2004; Acquisti et al., 2016; Sundar et al., 2013). Furthermore, abuse of information is often hidden and invisible to consumers, and ambiguously formulated privacy policies cause misinterpretation of data sharing practices. This situation eventually leads to false assumptions about privacy protection and enhances information asymmetry (K. Martin, 2013; Reidenberg et al., 2015). The principles of notice and choice require users to fully understand the terms and condition of the commercial agreement. However, in a digital marketplace where the costs of identifying and fixing a privacy breach are high—and uncertainties about information flow and information asymmetry run rampant—consumers are trapped in their bounded rationality, and these principles are likely to fail (Nissenbaum, 2011, Simon, 1972).

Based on the three aspects discussed above, it can be concluded that to satisfy the principles of notice and choice, consumers must be better informed about data handling practices (e.g., by means of visualizations that summarize a privacy policy) or that the sole responsibility for potential risks must be shifted away from the data subject (e.g., by means of corporate privacy responsibility; Bandara, 2020; K. Martin, 2013). However, as fully informed consent is almost impossible to achieve, shifting the responsibility of data exchange away from a

purely legal contract to a social contract governed by implicit rules developed within specific situations and based on moral values might overcome this limitation. By doing so, the economic relationship is balanced and beneficial for all parties. According to social contract theory, this allows consumers to (re)-gain power and to differentiate between relationships and to discriminately share information, reaching a balance between withholding and sharing information dependent on the context (K. Martin, 2016a, 2016b; K. Martin & Shilton, 2016).

Hence, an economic relationship associated with data exchange depends on interdependency between the involved parties. A violation of any former agreements and the usage of data for purposes other than declared results in an imbalance of economic power. Assessing the data market environment through the lens of the power–responsibility equilibrium framework (K. Davis et al., 1980), upon data exchange, the data recipient can be considered the power holder. The consumer that discloses information relies on the responsible treatment of the provided data (e.g., no sale of user data to third parties) and protection against misuse (e.g., application of suitable security measures). Violation or exploitation of this asymmetric power relationship between data subjects and data holders elicits privacy concerns, a sense of helplessness, anxiety, uncertainty, fear or defensive behavior such as termination of the trade relationship, or withholding information (Bandara, 2020; Bandara et al., 2019; Lwin et al., 2007; Nissenbaum, 2011).

Borrowed from traditional approaches to consumerism, the equalization of the market, characterized by a healthy and balanced relationship between companies and consumers, is the basis for a flourishing and fair economic environment (Kucuk, 2009). In a digital market environment, the power relationship can quickly become imbalanced due to the vast amounts of data in circulation and the seemingly insatiable appetite of consumers that opens the doors for new vulnerabilities and risks, affecting consumers' rights and information privacy (Kucuk, 2016). Although the foundations have been laid for integrating privacy protection into the digital market (e.g., GDPR or CCPA), the time has now come to enhance consumer empowerment, reducing risks and vulnerabilities for the consumer and paving the way to improved fair trade relations.

1.5 Data privacy zones, uncertainty and context: A social perspective on information privacy

“The opportunities for unwitting disclosure in cyberspace range from uses of electronic identification for access to the system, to..., identification of viewer preferences,...., and countless ways in which computerized identification is coupled with personal preferences and behavior. Americans [and users in general] have very little comprehension of or agreement about what privacy entails, but 84% are concerned about it...” (Branscomb, 1995, p. 164).

In the information age, many activities from the offline world, such as talking, dating, news consumption, shopping or healthcare, are now also possible via online platforms. The large-scale availability, fluent and fast-moving online environment requires individuals to constantly exchange information, to consider the appropriateness of information disclosure and to define their privacy preferences. Whereas the boundary between public and private is easier to define in the offline world, this distinction seems to diminish in an online environment. The proportion between the public and private sphere and revealing or withholding information depend not only on social norms that exist within a culture at a certain point in time but also on social relationships, status, life situations and the technology itself (Marx & Muschert, 2007; Ochs & Löw, 2012; Simmel, 1992; Trepte & Masur, 2020; Westin, 1967). Hence, new norms for privacy behavior are developing, determined by social structures and technology in the online world (Ochs & Löw, 2012; Trepte & Masur, 2020). Apart from privacy as a sociotechnical system (Nissenbaum, 2010), individual attitudes also shape privacy preferences.

Attitudes can be formed by a rational consideration of facts where pros and cons are carefully weighed against each other. However, in many cases—and particularly in situations characterized by uncertainty—attitudes are not formed rationally (Gleitman et al., 2011) but rather through (1) classical conditioning (through repeatedly pairing two stimuli; dating back to Pavlov, 1927), (2) operant conditioning (through rewards; dating back to Skinner, 1938), or (3) observational learning (through observing behavior of others; dating back to Bandura, 1969). Moreover, the role of trust and emotions in shaping privacy attitudes should not be underestimated. However, the centralized nature of internet services, the opaque structures of data highways, the asymmetry in power relations or the take-it-or-leave-it nature of many online services make trust management a challenging task for individuals operating in the cyberworld (Arkko, 2020; Berendt et al., 2005; Bräunlich et al., 2020; Kelbert et al., 2012; Müller et al.,

2012; Ochs & Löw, 2012). The problem with privacy is not that people do not care (Branscomb, 1994) but that most individuals lack knowledge about data handling practices, actual usage and the consequences of data disclosure (Acquisti, 2004; Schneier, 2015). Whereas most people can indicate and articulate what their physical privacy zones are, an intangible online environment leads to a situation in which people encounter difficulties grappling with their data privacy zones and determining privacy boundaries (Schwartz, 1999). The intangibility of privacy boundary violations online and the (near) impossibility of assessing the probability of breaches leaves individuals in a state of *uncertainty*. Uncertainty is mainly caused by four factors: (1) information asymmetry, (2) bounded rationality, (3) biases affecting decision-making and (4) the context (Acquisti, 2009; Acquisti et al., 2017).

First, one of the main factors that cause uncertainty toward privacy boundaries is information asymmetry. Information asymmetry is manifested by an imbalance of power between the data subject and data holder. Usually, data subjects are less informed about the value of their personal data and data processing practices than the data holder. This state of incomplete information prevents rational assessment of the data disclosure situation (Harsanyi, 1967). Such a situation of incomplete or uneven information distribution can arise through invisible technical processes running in the background (e.g., creation of databases or data mining), the opaque structure of data highways (e.g., where does the data actually go), the obscurity of the consequences of data disclosure (e.g., profiling is often intangible) and the likelihood of privacy violations (e.g., how likely is one to become a victim of identity theft; Acquisti, 2004; Acquisti et al., 2015; Bräunlich et al., 2020; Flender & Müller, 2012).

Second, users are bounded by human nature and do not possess the cognitive capabilities to process all relevant information (Simon, 1982). Even when information about data handling practices is readily available, the required cognitive involvement to derive a decision is either often disproportionately high in relation to the intended objective or individuals seal off due to information overload (Acquisti, 2004; Acquisti & Grossklags, 2005; Veltri & Ivchenko, 2017). To address cognitive limitations and information overload, information disclosure behavior is often guided by heuristic thinking—mental short cuts—rather than systematically assessed (Acquisti et al., 2017). Heuristics ‘assist’ individuals arrive at a decision quickly, even in uncertain situations. Relying on intuition often corresponds with a positive attitude toward information disclosure (Sundar et al., 2013), and future consequences of data disclosure are ignored, misjudged or underestimated (Acquisti, 2004; Jia et al., 2015).

Third, apart from the complexity of information and asymmetry of its distribution, human decision-making and risk-benefit assessments are also affected by biases. Biases are often subconscious, strong, systematic, mostly consistent over time and situations—and difficult to reinfluence or ignore. Often, the probability of serious consequences resulting from information disclosure is judged to be small, as people tend to be overly reliant on the fact that everything worked out fine before (Irwin, 1953). Optimistic bias tricks individuals to underestimate the risk probability caused by unsafe online behavior or to overestimate the ability to protect against privacy intrusion. This tendency is closely related to confirmation bias, as people are in favor of evidence consistent with previously made decisions or beliefs. Self-control bias (Loewenstein, 1999) and hyperbolic discounting bias (Acquisti & Fong, 2020) entice humans to favor immediate benefits or to enjoy the heat of the moment while suppressing or miscalculating future consequences. Furthermore, the disclosure or withholding of personal data is malleable and can be manipulated by external entities or design choices (framing bias). Framing is often subtle and unrecognizable for individuals, but it is a strong mechanism that has the power to trigger heuristic thinking. For instance, a message is framed in terms of loss/gain or using red/green colors to represent danger/safety (De Martino et al., 2006). The power of such nudging mechanisms can be used for good (improving privacy behavior; Garg & Camp, 2013) or bad (provoking data disclosure; Acquisti et al., 2017; Acquisti et al., 2015). Whereas ‘cookie notices’ are intended to assist informed choices, ‘malicious interfaces’ elicit the opposite effect (Conti & Sobiesk, 2010). For instance, opt in is usually the default setting, and opting out takes additional effort, which encourages users to blindly click through cookie notices rather than actively engage in making informed choices (Utz et al., 2019). Relying on mechanisms borrowed from cognitive psychology, those dark patterns are used to implement features into the interface design that deceive users into taking actions that are not in their best interest. This manipulative design practice can impact users’ privacy or limit the execution of privacy protection (Gray et al., 2018; Mathur et al., 2019; Waldman, 2020). Dark patterns are ethically questionable, as they draw on cognitive mechanisms and biases that are executed automatically, usually go unnoticed and are very hard to inhibit.

Fourth, online behavior is not driven by psychological factors alone. The context also influences privacy preferences and information disclosure. Individuals can value their privacy on a general level, but as soon as a concrete transaction takes place, previous privacy valuation seems to be relativized. Here, instead of potential risks and benefits being considered, context and situational factors predominately guide the decision-making process (Müller et

al., 2012). The theory of contextual integrity (Nissenbaum 2004, 2011) assumes that information flow is amenable to social boundaries and contextual norms. To overcome limitations in human decision-making and the near impossibility of giving truly informed consent, information exchange and a potential violation of information privacy must be considered in light of context-specific informational norms. These norms are defined by the actors involved, the nature of the information (information types), and in terms of appropriateness and information distribution (transmission principles). Additionally, unstated social norms based on moral values also play a role in information exchange and allow users to discriminately share information and to decide when information disclosure is beneficial or damaging (social contract theory; Acquisti et al., 2016; K. Martin, 2016a, 2016b; K. Martin & Shilton, 2016).

From the aforementioned discussion, it can be inferred that uncertainty due to *incomplete information*, cognitive limitations resulting from *heuristic thinking* and subconscious *biases* and the *context dependency* of privacy can lead to a situation where actual behavior is often contrary to stated intentions (Acquisti, 2004; Acquisti et al., 2015; Norberg et al., 2007). While users theoretically—and on a general level—indicate interest in their online privacy and show willingness to protect or limit data disclosure, this is often not reflected in actual behavior. This contradictory behavior toward online privacy, also known as the privacy paradox, is heavily discussed in the scientific literature. Some scholars argue that the reason for this indifference lies in our busy lifestyles or in the unfamiliarity with (future) risks (A. L. Allen, 2012). Others seek answers regarding cognitive limitations (Deuker, 2010), habits (Debatin et al., 2009), emotions that implicitly guide decision-making (Schwarz, 2012), the need for interpersonal relationships (Debatin et al., 2009), seemingly free services, the sharing economy, or the perceived need to improve one's life. Moreover, repeated privacy intrusion can lead to a state of helplessness and ignorance of privacy concerns (Shklovski et al., 2014).

To understand what privacy actually means, Solove (2002) favors a pragmatic approach for conceptualizing privacy. This essentially means approaching privacy from the bottom up under consideration of the context instead of further relying on an abstract conceptualization that still tries to separate the public from the private space. The key for understanding privacy would then be to define practices that might disrupt privacy, for instance breach of confidentiality, loss of control over information, or surveillance. Through the lens of contextual integrity, the three parameters that define contextual norms (actors, information types, and transmission principles) are independent and not counterbalanced but complementary to each other. It is therefore impossible to reduce privacy to

one parameter only. As technology and social structures change over time—as the significance of family, home and body has evolved over time—the practices associated with privacy also change. Hence, the information consciously or unconsciously disclosed is shaped by sociotechnical developments, individual attitudes, psychological factors and context. To understand the value of privacy, we must understand “the problem itself than trying to fit the problem into a general category” (Solove, 2002; p. 1154).

1.6 Research goals and question

“Privacy is not the opposite of sharing — rather, it is control over sharing”
(Acquisti et al., 2016, p. 445).

Compared to traditional media, the internet and its associated technologies are probably the most significant innovation of our time. As with many aspects in human lives, technological development is also a double-edged sword. On the one hand, individuals appreciate the benefits of online services. On the other hand, they worry about the tremendous and irreversible consequences that come with those technologies (Buchman, 2012). Raising the question: ‘To what extent does technology influence privacy’ would arguably provoke the need for a wide range of answers.

A historian would probably argue that technological development, over time, heavily shaped humans’ perception toward privacy. While autonomy and the privacy that comes with it only applied to the privileged in former times, privacy has become a universal human right for citizens living in a democratic society since the middle of the 20th century. While the right to privacy is easier to define in the offline world, it becomes increasingly opaque in an online environment. Sophisticated technologies leave individuals with concerns about their right to privacy but have also produced advanced online societies that previously never existed.

Many legislators would probably agree that the law is continually playing ‘catch-up’ with technological developments and that current regulations are stopgaps rather than viable long-term online privacy protection measures. However, a legislator would also emphasize—and deservedly so—that much has happened in the field of privacy protection law. The European GDPR and the California CCPA came into force in 2018 and 2020, respectively, demonstrating that legally binding privacy protection regulations are essential if users are to be given autonomy and control over their personal data.

A technician would possibly argue that technology can either be used as privacy infringing or as privacy enhancing. From a technical perspective, privacy has to do with the collection, usage, and sharing of data, and its protection relies on cybersecurity mechanisms such as cryptography, authentication and authorization. As giving truly informed consent is almost impossible to achieve, the scales between the two sides of what technology can be used for must be tipped toward privacy-enhancing technologies.

An economist might stress that data clearly have value and that individuals and societies can benefit from new technologies and data analysis techniques. The current COVID-19 crisis impressively demonstrated the benefits of data. However, an economist would also look at the potential power imbalances between data subjects and data holders.

A social scientist would probably admit that it is still unclear which factors exactly guide privacy behavior but that the determination of privacy preferences is an interaction between several internal and external factors, dependent on context and not rationally calculated. Limitations to human decision-making, information asymmetry, bounded rationality, the occurrence of biases and context dependency indicate that individuals are unable to engage in a mere cost-to-benefit analysis.

The discussion of information privacy from various perspectives shows that as technologies become increasingly complex, so too does the realm of what is considered private information. Although much progress has been made in the development of legislation pertaining to privacy protection, it seems that existing regulations, such as the GDPR, are broadly defined and still in their infancy (Perry, 2019; Zarsky, 2017). Moreover, it is reasonable to assume that educating the user to gain a thorough understanding of all the technical processes is challenging and technically knowledgeable users will be the exception rather than the rule. Furthermore, the power relationship will most likely remain unbalanced, with scales tipping toward large market players such as Google, Apple and Facebook (Craig & Ludloff, 2011). Moreover, the cognitive limitations influencing decision-making, such as biases and heuristic thinking, are very difficult to overcome (Novak & Hoffman, 2008).

To counteract these difficulties and to get closer to the provision of *notice*, *choice* and *consent*, it is important to support the user. Providing information about the data handling practices of online services and communicating the potential risks to users seem to be a viable solution. However, this must be done in a ‘user-friendly’ fashion. This means providing information in a visually appealing and easy-to-access/understand manner. By doing so, users are provided with a reasonable chance to consciously make decisions about their online privacy.

Privacy can be considered a prerequisite for autonomy. To form self-determined opinions, users must have the freedom to decide for themselves what information they want to share, with whom and for what purpose. New technologies place a heavy burden on users to safeguard this freedom, as the online landscape makes it nearly impossible to control data flow. However, it is not the existence of privacy that is up for debate, rather its importance, scope and deployment.

In line with this, this dissertation attempts to gain insights into the online privacy behavior of users by mainly focusing on the social perspective—combined with insights from the technical perspective—on information privacy. The aim of this dissertation is not to judge technological innovations as bad or good but to understand “the coexistence of acceptance and fear with regard to internet usage” (Buchmann, 2012; preface) and to answer the overall research question:

How can users of online services be empowered to protect their online privacy while balancing the scales between privacy preferences and the potential risks associated with information disclosure?

This dissertation addresses two research goals to help answer this question. The first goal centers around knowledge acquisition, and the second research goal focuses on a design perspective:

1. *To gain insights into the online privacy behavior of users.*
 - a. To assemble a comprehensive overview of theoretical explanations for the privacy paradox.
 - b. To explore the role of technical knowledge and privacy awareness on users’ online behavior.
 - c. To gain insights into differences in users’ privacy perceptions and preferences.

2. To design a research-based privacy awareness tool to empower users to make well-informed decisions about their online privacy.

- a. To derive a comprehensive understanding of generally applicable privacy attributes of online services.
- b. To develop and evaluate a user-centered privacy visualization.

To achieve these research goals, a multimethod approach is used that combines knowledge derived from systematic literature reviews, qualitative and quantitative research studies and user experience (UX) design. Moreover, the research goals are approached using an interdisciplinary approach that combines the social and technical perspective and considers existing privacy guidelines.

1.7 Dissertation outline

To gain insights into the online privacy behavior of users, four studies were conducted. First, a systematic literature review of theories explaining the privacy paradox is presented in **Chapter 2**. Recent research on online behavior has revealed discrepancies between users' privacy attitudes and their actual online behavior. While users claim to be concerned about their privacy, they nevertheless undertake very little to protect their personal data. The systematic analysis of the existing literature aims to address seemingly paradoxical behavior through different theoretical lenses, contributing to the first research goal of this dissertation (1a). Furthermore, this literature review serves as input for the three empirical studies conducted to help better understand the consideration-process users implement when using online services and their online privacy preferences.

Chapter 3 describes an experiment that tests the privacy paradox and aims to address deficiencies in the current privacy paradox literature. First, research focusing on the privacy paradox in a mobile environment is still limited. Second, the actual behavior of users is seldom measured in regard to seemingly paradoxical behavior online. Third and fourth, it is often assumed that a lack of knowledge and awareness can lead to paradoxical behavior, as well as financial considerations. To eliminate the effects of a lack of technical knowledge and privacy awareness, an experiment on the downloading and usage of a mobile phone app among 66 technically savvy students was designed, giving them sufficient money to buy an app. This study contributes to the first research goal of this dissertation (1b).

In **Chapter 4**, an interview study with experts is presented. Most insights into the privacy paradox are based on research among general users. It is unclear whether users with high expertise pertaining to online privacy and cybersecurity would show similar discrepancies between concerns and behavior. To gain insight into this question, 20 privacy and cybersecurity experts were surveyed about their views on online privacy regarding mobile apps, contributing to the first research goal of this dissertation (1b).

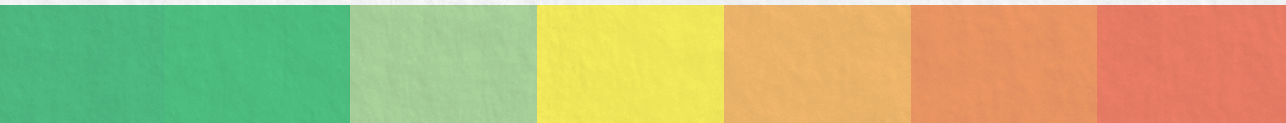
Chapter 5 presents a Q-sort study that was conducted with 100 general users of online services. This study was designed to gain insights into the context dependency of online privacy. To provide a deeper understanding of users' online privacy perceptions, this study considered privacy in two contexts (health vs news app) and focused on three contextual factors: what information is processed, how it is processed, and who processes it. This study contributes to the first research goal of this dissertation (1c).

To gain insights into the question of how to visually communicate to users the most relevant aspects of a privacy policy, two studies were conducted. **Chapter 6** presents a further systematic literature review. To gain knowledge about privacy attributes playing a role in online services, existing Privacy by Design guidelines and proposals of privacy visualization were reviewed. Based on the reviewed guidelines and proposals from academia, industry and government, a unified list of privacy attributes was distilled and ranked according to perceived importance, as indicated by 385 users and 100 privacy experts. This unified list of privacy attributes provides a foundation for user-centric privacy visualizations and contributes to the second research goal of this dissertation (2a).

Eventually, the knowledge derived from the literature reviews and empirical studies was combined into a proposal for visualizing data handling practices. Many countries now mandate transparency and consent when personal data are handled. However, most users do not read privacy policies or are not able to fully understand them. Privacy visualizations can alleviate this problem, but existing approaches are incomplete and not user centered.

Chapter 7 describes the design of a privacy visualization, called the *Privacy Rating*. Furthermore, a usability test with 20 users aimed at evaluating the visualization is presented. This chapter contributes to the second research goal of this dissertation (2b).

The dissertation concludes with **Chapter 8**, which presents a discussion of the results and the theoretical and practical implications derived from the conducted research—aiming at answering the main research question of this dissertation.



2

The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review

Barth, S., & De Jong, M. D. T. (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics*, 34, 1038-1058. <http://dx.doi.org/10.1016/j.tele.2017.04.013>

2.1 Introduction

The emergence of the Semantic Web has brought numerous opportunities with it, including an almost unlimited access to information, round-the-clock social networking connectivity and large scale data aggregation. It has grown to the extent that it now plays a part in the everyday lives of billions of people around the world. Simultaneously, the advent of big data and digital technologies have also raised serious privacy and security issues. E. J. Smith and Kollars (2015) called these digital developments “the uncontrolled electronic panopticism” (p. 160). Fact is, the information being transformed between electronic devices equates to a form of unwitting user observation. When considering mobile applications and data ‘leakage’ in particular, recent literature argues that the consumer’s choice to use mobile technologies is primarily driven by considerations of popularity, usability and the price of a given technology (Kelley et al., 2013; Kim et al., 2008) despite the potential risk of data misuse. At the same time however, research indicates that consumers are concerned about their privacy, including the ambiguous distribution of data and their use by third parties (H. J. Smith et al., 2011). This discrepancy between the expressed concern and the actual behavior of users is a phenomenon known as the privacy paradox: users claim to be very concerned about their privacy but do very little to protect their personal data.

There are currently multiple theories explaining the privacy paradox. Some have explained this paradoxical behavior from a rational perspective by arguing that users weigh the cost-benefit ratio of online information disclosure both consciously and rationally (Simon, 1955). Others have questioned this rational view by arguing that individuals are bound in their rational decision-making by several cognitive biases, resulting in a pre-determinable cost-benefit calculation (Simon, 1982). Interestingly, both perspectives result in a risk-benefit calculation that ultimately chooses benefits over risks. In addition, an unbalanced decision-making process serves as the basis for a third perspective, where decision-making is based on prevalent benefits and as a result, no or negligible risk assessment takes place.

Before introducing the present systematic literature review, the phenomenon of the privacy paradox will be discussed. After introducing the methodology, a review of the different theoretical approaches to the phenomenon will be presented. Lastly, the results will be discussed in terms of the nature of decision-making, the context within which the disclosure behavior takes places and solution-oriented implications.

2.1.1 The privacy paradox

The majority of research into the privacy paradox considers general internet activities with a focus on e-commerce and social networking activities in particular. Known as the privacy paradox, it is a documented fact that users have a tendency toward privacy-compromising behavior online which eventually results in a dichotomy between privacy attitudes and actual behavior (Acquisti, 2004; Barnes, 2006). A certain degree of risk perception implies greater knowledge of privacy protection strategies but appears an insufficient motivator to apply such strategies (Oomen & Leenes, 2008). Thus, while many users show theoretical interest in their privacy and maintain a positive attitude toward privacy-protection behavior, this rarely translates into actual protective behavior (Joinson et al., 2010; Pötzsch, 2009; Tsai et al., 2006). Furthermore, while an intention to limit data disclosure exists, actual disclosure often significantly exceeds intention (Norberg et al., 2007).

Research into online service providers has shown that concrete privacy decisions and abstract risk awareness are not interchangeable. Privacy decisions do not change in line with modified preferences, which could explain the disparity between stated preferences for privacy and actual behavior (Flender & Müller, 2012). Although users are aware of privacy risks on the internet, they tend to share private information in exchange for retail value and personalized services (Acquisti & Grossklags, 2005; Sundar et al., 2013).

In the context of users' social network activities, a similar pattern is observed. The utilization of various privacy protection strategies such as limiting wall post access, restricting photo tags and sending private messages instead of posting open content is designed to control the flow of information between friends and peers. Implementing such strategies however, shows little concern for data collection by third parties in the background (A. L. Young & Quan-Haase, 2013). Privacy concerns should logically lead to restricted provision of information in social networks; however, the reverse effect can be observed as many users provide personal information seemingly without any hesitation (Hughes-Roberts, 2012; Manier & O'Brien Louch, 2010; Nagy & Pecho, 2009; Yoo et al., 2012).

Looking at the provision of personal information in an app purchase process, Buck, Horbel, Germelmann, et al. (2014) found that information from one's social group and the app store itself is more relevant than actual information about exploitation of personal data by third parties. Users are able to articulate their privacy needs but the actual decision to use (context-aware) applications does not align with their claims. Oetzel and Gonja (2011) go a step further, stating: privacy is not yet integrated into the social presentation of a smartphone and hence, will consequently lead to failed privacy awareness. Thus, supporting privacy

awareness with suitable tools would allow meaningful decision-making by users and solve the conflicting interests between users and providers (Deuker, 2010).

The previous discussion showed a variety of views on the emergence and existence of the privacy paradox. However, the cited literature discusses the discrepancy between privacy concerns and actual information disclosure from a practical point of view as it can be observed within the context of general internet activities, e-commerce, social networking sites and mobile applications. But what does the theory say about this phenomenon and why do users take so many risks?

To our knowledge, there is not one unilaterally accepted theory used to explain the online behavior of users when it comes to information disclosure, nor is there a consensus on the mental processes users rely upon when deciding whether to disclose information or not. We did however, find a review of current research on the privacy paradox phenomenon (Kokolakis, 2017). The paper deals mainly with literature that either supports or challenges the existence of the privacy paradox under consideration of a number of theories, stressing the need for one theoretical model. The paper does not provide a full theoretical discussion of the phenomenon nor does the author offer any new ideas for solving the privacy paradox.

Our systematic literature review on the other hand, attempts to develop an overarching theoretical framework, addressing the discrepancy between privacy concerns and actual online protective behavior through different theoretical lenses with a special focus on mobile applications.

2.2 Method

This chapter presents a systematic literature review of all the studies that discuss the phenomenon of the so-called privacy paradox in the online environment. The main focus will be on mobile applications but as only nine studies addressing the subject could be ascertained via a literature search, the parameters in which the privacy paradox might be relevant were broadened to include social network sites, general online contexts, websites and e-commerce platforms.

An electronic database literature search was conducted in GoogleScholar, Scopus, IEEE, Web of Science and ACM. The keyword 'privacy paradox' was used as the primary broad search string. Papers were selected by their relevance as indicated by title or abstract and a subsequent examination of the full paper. Furthermore, a manual search of reference lists was conducted to identify additional papers that may have been missed by the electronic database search. Overall, only full, peer-reviewed papers, peer-reviewed conference papers and

published book-chapters were included. Only the subject articles in which the phenomenon of privacy paradox was explicitly mentioned and discussed were considered eligible. Articles that dealt with privacy issues and concerns in general were not included. These searches resulted in an item set of 110 articles.

In a second step, the articles were analyzed according to theories that had been applied to approach the phenomenon of the privacy paradox. Only articles that discussed the privacy paradox explicitly with the aid of a theory or theoretical concept were included in the final sample. Hence, articles that applied no theory at all or a theoretical approach or umbrella terms (“social capital”, for instance) for discussing the phenomenon were excluded. Furthermore, articles discussing the personalization privacy paradox (a special subcategory of the privacy paradox) or solely providing a practical solution perspective, legal and ethical discussions/papers, research proposals and commentaries were excluded from the sample. This eventually resulted in a final sample of $N = 32$ full papers that accounted for 35 theories in total.

In a third step, the articles were reviewed in detail with the aim of detecting clusters or divisions into which the different theories could be assigned. A pattern emerged, making a differentiation between rational and irrational approaches for paradoxical behavior at an intrapersonal or interpersonal level possible. Furthermore, the majority of theories centered around a given cost-benefit calculation, with most favoring benefits over costs. An overview of all the clusters and the corresponding theories and articles discussed in the next section of this paper can be found in Table 2.1.

2.2.1 Approaches to the privacy paradox

When examining the nature and factors of decision-making, the applied theories ($N = 35$) were clustered into two main categories: (1) decision-making based on a risk-benefit calculation or (2) decision-making based on prevalent benefits and little to no risk assessment. The risk-benefit calculation category can be further divided into (1a) a largely rational calculation of risks and benefits whereas benefits outweigh risks, and (1b) a calculation process but risk assessment is biased and benefits are prevalent as a result. Both calculation processes ultimately lead to paradoxical behavior eventually. Category (1b) can be further divided into five different types of biases influencing the calculation process: (I) heuristics, (II) under- and/or overestimation of risks and benefits, (III) (immediate) gratifications, (IV) difference between judgments of risks and benefits, and (V) habit. The main category where no or marginal risk assessment only takes place accounts for three sub-categories: (2a) value of desired goals outweighs risks assessed. Additionally, this process is largely determined by

in- and out-group considerations, (2b) the privacy valuation failed, and (2c) a knowledge deficiency eventuates as a result of incomplete information. The latter sub-category is also part of the bounded rationality approach but because it is deemed a failed risk assessment due to incomplete information, it was assigned to the second main category. In the following paragraph, the mentioned categories will be discussed. For an overview of the categorization of the various theories, please see Figure 2.1. A comprehensive description of the theories can be found in Appendix 2.1.

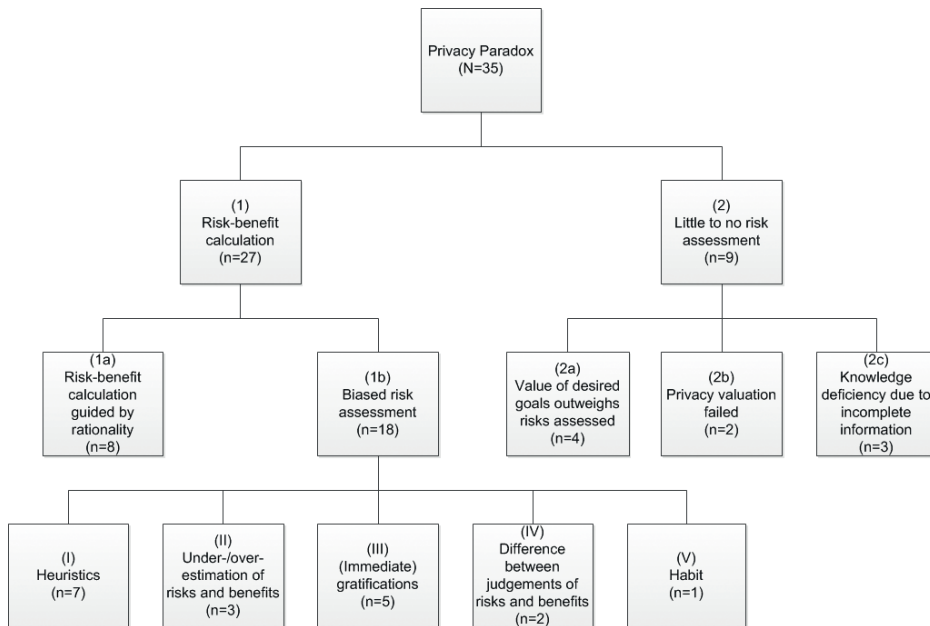


Figure 2.1 Overview of categorization theories according to nature of decision-making

Table 2.1 Overview of theories according to main categories and sub-clusters and corresponding articles

Main category	Sub-cluster I	Sub-cluster II	Theory	Article(s)	Domain
Risk-Benefit Calculation	Risk-benefit calculation guided by rationality	-	Rational Choice Theory of Human Behavior (Simon, 1955)	Hu and Ma (2010)	SNS
			Adaptive Cognition Theory of Social Network Participation (Hu & Ma, 2010)	Hu and Ma (2010)	SNS
			Privacy Calculus Theory (Culnan & Armstrong, 1999)	H.-T. Chen and W. Chen (2015); Dinev and Hart, (2006); Kehr et al. (2014); Motiwalla et al. (2014); Pentina et al. (2016); Poikela et al. (2015); Wilson and Valacich (2012)	e-commerce; mobile application; SNS; website
			Resource Exchange Theory (Donnenwerth & Foa, 1974; Foa, 1971)	Wilson and Valacich (2012)	e-commerce
			Expectancy Theory (Vroom, 1964)	Wilson and Valacich (2012)	e-commerce
			Rational Ignorance Theory (Downs, 1957)	Flender and Müller (2012)	social web
			Theory of Reasoned Action/Theory of Planned Behavior (Ajzen & Fishbein, 1980) Ajzen 1985)	Poikela et al. (2015)	e-commerce; SNS

Table 2.1 *Continued.*

Main category	Sub-cluster I	Sub-cluster II	Theory	Article(s)	Domain
			Dual Process Model of Cognition (System II) (Kahneman, 2003)	Phelan et al. (2016)	general online context
	Biased risk assessment within the risk-benefit calculation	Heuristics	Theory of Bounded Rationality (Simon, 1982)	Acquisti (2004); Acquisti and Grossklags (2005); Deuker (2010); Flender and Müller (2012); Jia et al. (2015); Pötzsch (2009)	e-commerce; mobile apps; SNS; social web
			Cognitive Heuristics (Tversky & Kahneman, 1975)	Gambino et al., (2016); Sundar et al. (2013); Wakefield (2013)	e-commerce, mobile websites
			Extension to the Privacy calculus theory (Culnan & Armstrong, 1999)	Kehr et al. (2015)	mobile application
			Cues-filtered-out Theory (Sproull & Kiesler, 1986, 1991)	Pötzsch et al. (2010)	web forums
			Feelings-as-Information Theory (Schwarz, 1990, 2012)	Kehr et al. (2014)	website
			Structuration Theory (Giddens, 1984)	Zafeiropoulou et al. (2013)	mobile application
			Communication Privacy Management Theory (Petronio, 1991, 2002)	Sundar et al. (2013)	e-commerce; SNS

Table 2.1 *Continued.*

Main category	Sub-cluster I	Sub-cluster II	Theory	Article(s)	Domain
		Under-/overestimation of risks and benefits	Optimistic Bias Theory (Irwin, 1953)	Acquisti (2004); Flender and Müller (2012)	e-commerce; social web
			Theory of Under Insurance (Kunreuther, 1984)	Acquisti (2004)	e-commerce
			Third-Person Effect Theory (Davison, 1983)	Debatin et al. (2009)	SNS
	(Immediate) gratifications		Immediate Gratifications (O'Donoghue & Rabin, 2001)	Deuker (2010); Flender and Müller (2012); Acquisti (2004); Wilson and Valacich (2012)	e-commerce; mobile application; social web
			Self-Control Bias (Loewenstein, 1999)	Acquisti (2004); Acquisti and Grossklags (2005)	e-commerce
			Hyperbolic Discounting Theory (Laibson, 1997)	Acquisti (2004); Acquisti and Grossklags (2005); Flender and Müller (2012); Hughes-Roberts (2013); Wilson and Valacich (2012)	e-commerce; SNS; social web
			Theory of Cognitive Absorption (Agarwal & Karahanna, 2000; Agarwal et al., 1997)	Alashoor and Baskerville (2015)	SNS

Table 2.1 *Continued.*

Main category	Sub-cluster I	Sub-cluster II	Theory	Article(s)	Domain
			Uses and Gratification Theory (Blumler & Katz, 1974; Katz et al., 1974)	Debatin et al. (2009); Quinn (2016)	mobile websites; SNS
		Difference between the judgements of risks and benefits	Prospect Theory (Kahneman & Tversky, 1979)	Hughes-Roberts (2013)	privacy policies; SNS; website
			Quantum Theory (Based on Busemeyer et al., 2006)	Flender and Müller (2012)	e-commerce
		Habit	Theory of Ritualized Media Use (Rubin, 1984)	Debatin et al. (2009)	SNS
Little to no risk assessment	Value of desired goal outweighs risk assessment	-	Privacy Regulation Theory (Altman, 1975)	Shklovski et al. (2014)	SNS; social web
			Conformity and Peer group pressure (Crutchfield, 1955)	Flender and Müller (2012)	social web
			Duality of Gemeinschaft und Gesellschaft (Tönnies, 2012)	Lutz and Strathoff (2011)	SNS
			Extended Two-Component Model of Self-Presentation Online (based on Leary & Kowalski, 1990)	Krämer and Haferkamp (2011)	SNS

Table 2.1 *Continued.*

Main category	Sub-cluster I	Sub-cluster II	Theory	Article(s)	Domain
	Privacy valuation failed	-	Public Value Theory (Meynhardt, 2009)	Lutz and Strathoff (2011)	SNS
			Social Representation Perspective (Moscovici, 1984; Abric, 1996)	Oetzel and Gonja (2011)	SN, Google; smartphone
	Knowledge deficiency due to incomplete information	-	Theory of Incomplete Information (Harsanyi, 1967)	Acquisti (2004); Acquisti and Grossklags (2005); Buck, Horbel, Germelmann, et al. (2014); Deuker (2010); Flender and Müller (2012)	e-commerce; mobile application; social web
			Dual Process Model of Cognition (System I) (Kahneman, 2003)	Phelan et al. (2016)	general online context
			Symbolic Interactionism (Blumer, 1986)	A. L. Young and Quan-Haase (2013)	SNS

2.3 Risk-benefit calculation guiding decision-making processes

Some of the studies claim that rational processes account for this paradoxical behavior, as decisions are carefully considered by way of conscious-analytic, profit-loss calculations. In other words: users consciously weigh the disadvantages of privacy disclosure against the perceived benefits. It would seem that users consciously resolve discrepancies between the willingness to obtain and possess something (such as downloading an app) and the simultaneous difficulties that arise in terms of unknown threats or risks (such as potential data usage by third parties). In other cases, the risk-benefit assessment is not completely rationally calculated, rather subject to and influenced by factors such as time constraints, immediate gratification or optimistic bias. Quite often, users are not consciously aware of such a bias and as a consequence, choose benefits while ignoring accompanied risks.

2.3.1 Risk-benefit calculation guided by rationality

Risk-benefit calculation plays a major role in the context of information privacy, known as the freedom to decide with whom, how and to what extent personal data are shared (H. Li et al., 2010). The cognitive style of decision-making during risk-benefit calculations is both analytical and conscious and can be described as “logical, cause and effect, rule-based, hierarchical, sequential, process-oriented, slower to implement but quicker to change, high effort, oriented toward delayed action, conscious, and experienced actively with the individual aware of and in control of the process” (Novak & Hoffman, 2008, p. 57). During information exchange, negative consequences are rationally weighed against goals and possible outcomes, aiming to maximize benefits and minimize the risks of information disclosure (Keith et al., 2013; Y. Li, 2012; Vroom, 1964). Hence, intention and actual behavior are positively influenced by expected benefits (e.g., using an app) but also negatively affected by associated costs (e.g., possible data usage by third parties; Culnan & Armstrong, 1999). Following, the different theories applied to the decision-making processes within the privacy calculation will be discussed. A comprehensive overview of all the variables that were mentioned in the following theories that play a role in the risk-benefit calculation is presented in Figure 2.2. According to *Rational Choice Theory of Human Behavior* (Simon, 1955) decisions are always reasonable and logical in order to gain the greatest benefit or satisfaction in line with an individuals’ perceived self-interest. In decision-making, individuals seek to maximize utility and minimize risk through rational calculus in response to both internal and

external constraints. When dealing with social media, users base their decision-making as it pertains to information disclosure on perceived benefits (e.g., networking with friends and acquaintances) and perceived risks (e.g., privacy and identity theft, image damage).

Building on this rational view of decision-making, Hu and Ma (2010) propose the *Adaptive Cognition Theory of Social Network Participation*. Here, user participation in online social networks can be assigned to three phases: initial use, exploratory use and managed use. The progression from one phase to the next results from understanding the benefits and risks associated, as well as the adaptation of activities and controls. The final phase can be described as an equilibrium of benefits and risk awareness formed by a continuous process of risk-benefit calculation.

Quite often, the perceived benefits outweigh the perceived risks, which eventually leads to the neglecting of privacy concerns that often results in the disclosure of information in exchange for social or economic benefit (*Privacy Calculus Theory*; Culnan & Armstrong, 1999). Within this calculation, economic benefits, personalization or convenience and social benefit tend to negate the downside of perceived risks (Wilson & Valacich, 2012). Individuals tend to concentrate on actual benefits rather than on previously stated concerns and calculated future risks with regard to issues such as location tracking by service providers via location-based mobile applications (Poikela et al., 2015). Moreover, an interdependency between the risk and benefit calculus exists as benefit valuations guide risk perception even if there is no relation in reality. Although users of social networks are confident they have taken adequate steps to control the flow of their private information (e.g., limiting profile visibility), this does not necessarily represent a significant decrease in the disclosure of personal information. This suggests that self-efficacy in privacy management on social networking sites can outweigh privacy concerns especially for those with low prior privacy concerns (H.-T. Chen & W. Chen, 2015). Furthermore, the cumulative effects of internet trust and personal internet interests can outweigh privacy risk perception to point that it eventually leads to the disclosure of personal information as it pertains to e-commerce transactions for example (Dinev & Hart, 2006). Pentina et al. (2016) added the Big Five personality factors and the influence of cross-cultural differences to the privacy calculus model. Extraversion and agreeableness increased perceived benefits from the use of mobile applications, regardless of the cultural environment. The satisfaction of informational and social needs led to the continued use of mobile applications despite the knowledge personal information might well be compromised. Privacy concerns did not have any influence on the adoption and

use of mobile applications. Even more, situational privacy calculus influences privacy disclosure as pre-existing attitudes (e.g., privacy concerns) may be fully overridden by situational factors such as the benefits associated with using a particular app (Kehr et al., 2014). Prior information disclosure behavior is generally more indicative of privacy valuation traits or individual privacy concerns. Fundamentalists for example (those highly concerned about data sharing with third parties and unwilling to place ongoing trust in particular organizations) expressed higher privacy valuation than those with a more pragmatic stance (concerned about data-sharing practices but willing to engage in privacy calculus activities), or unconcerned individuals (willing to share data, also with third parties without significant privacy concerns; Motiwalla et al., 2014).

In exchange for other resources such as money, services, time, status and love, people are willing to provide personal resources, including those in the form of personal information (*Resource Exchange Theory*; Donnerwerth & Foa, 1974; Foa, 1971). In such instances, personalization, convenience, economic benefits and social advantages will suppress the perception of risks while over-emphasizing the perceived benefits of privacy disclosure (Wilson & Valacich, 2012). Based on a subjective belief system, an individual chooses a certain behavior over another because of the expected outcome of maximizing benefits while minimizing costs (*Expectancy Theory*; Vroom, 1964). Here, the decision-making process is based on three beliefs: (1) valence (emotional attitude toward a particular outcome and the allure of receiving a reward); (2) expectancy (self-confidence to do s.th.); and (3) instrumentality (perception of the probability of gaining reward). As such, the conscious choice of an individual to ignore a certain piece of information is again based on a cost-benefit calculation, especially those where the informative effort (costs) are considered disproportionate to the perceived potential benefits (*Rational Ignorance Theory*; Downs, 1957). For instance, users may consider the cost of reading complex privacy policies in their entirety (e.g., loss of time or cognitive effort) outweighs the dangers, deciding that the benefits of using a service outweighs any potential privacy abuse concerns (Flender & Müller, 2012).

Taking the intentional perspective of a given rational behavior into account, the *Theory of Reasoned Action* (Ajzen & Fishbein, 1980) states that an individual's behavioral intention depends on the attitude toward a certain behavior and the subjective norms. The stronger the intention, the more likely it is that a person will engage in a certain behavior. The *Theory of Planned Behavior* (TPB) (Ajzen, 1985) takes this a step further as behavioral intention is influenced by existing attitudes about desired outcomes, social norms and the evaluation of the risk-benefit of that outcome (perceived behavioral control).

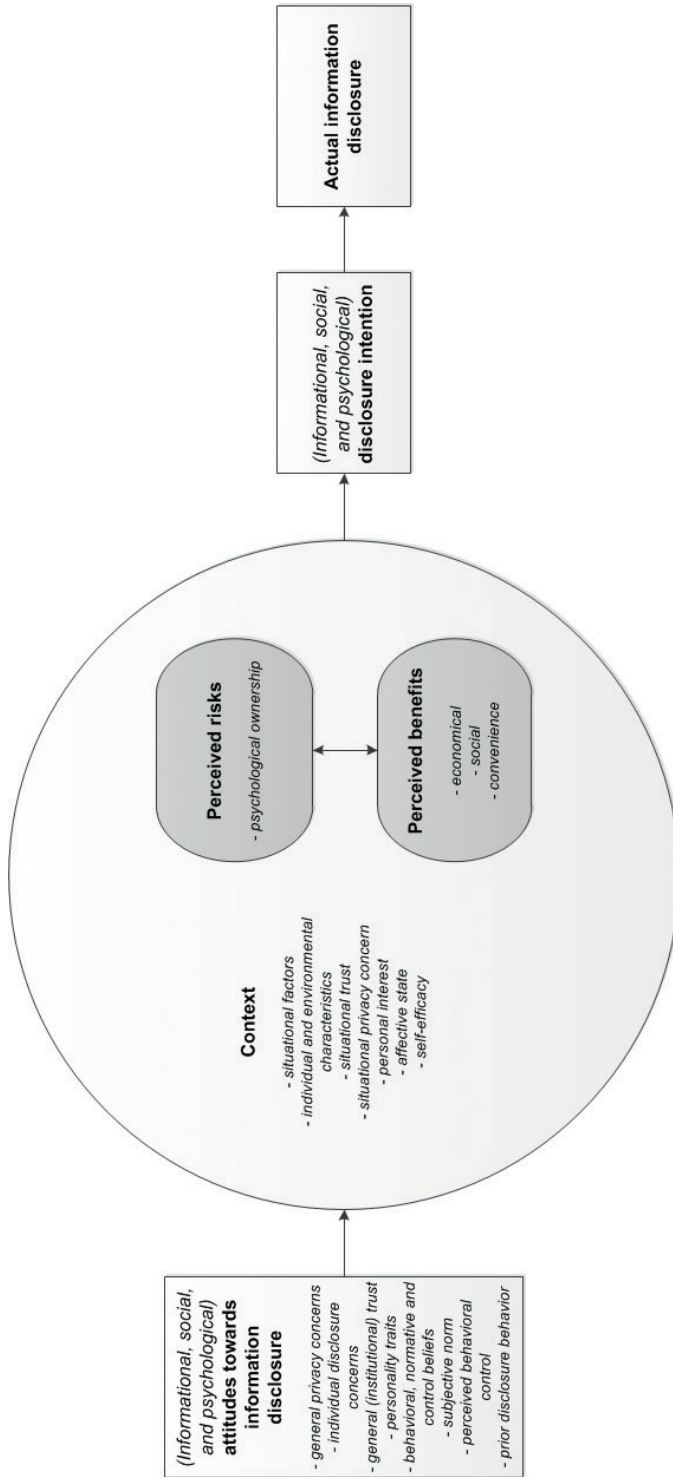


Figure 2.2 Overview of variables that play a role in the risk-benefit calculation guided by rationality

The stronger such attitudes, perceived control and the compliance with social norms are, the more likely it is that an individual will engage in a certain behavior. Actual control may be the missing variable that could explain the discrepancy between intention and behavior. The technological context of mobile apps adds various factors to the privacy calculus that are beyond the user's control such as near continuous tracking (Poikela et al., 2015). According to *Dual Process Model of Cognition* (Kahneman, 2003), decision-making is based on two systems. System I is fast and intuitive, whereas System II is rational and responsible for reasoning (resulting in legitimate concerns, for example). Intense intuitive concern may be overruled (but not reduced) by less intense concerns resulting from conscious consideration. Thus, privacy concerns are considered but most individuals are unable to address them adequately (Phelan et al., 2016).

The previously discussed theories show that the decision process as it pertains to information disclosure is guided by rational cost-benefit calculations where benefits outweigh risks. However, the majority of the analyzed studies showed a markedly different tendency as the theories to explain the privacy paradox can be characterized by the nonrational processes of decision-making. Here, the risk assessment within the risk-benefit calculation is biased by internal or external forces.

2.3.2 Biased risk assessment within the risk-benefit calculation

Contrary to a risk-benefit calculation guided by rationality, decision-making can also be influenced by different kinds of biases such as time constraints, time inconsistency, immediate gratification and optimistic bias. These biases are often non-conscious but play a major part in the eventual decision-making. Furthermore, bounded rationality also has an influence on the decision-making process. Too many options, unpredictable consequences, uncertainties or cognitive limitations eventually lead to subconscious biases in the calculation process. Hence, a decision is usually rapidly derived without an orderly subsequent analysis of the situation. As such, it cannot be verified according to process or grounds for judgment, but is instead based on experience (formal knowledge or beliefs) or a confidence in virtue. The experiential processes are “associative, emotional, low effort, rapid to implement but slow to change, parallel, immediate, outcome-oriented, holistic, preconscious and experienced passively with the process opaque to the individual” (Novak & Hoffman, 2008, p. 57). Following, the different theories of biased risk assessment within privacy calculation will be discussed. A comprehensive overview of all the variables that were mentioned in the theories that play a role in biased risk-benefit calculation is presented in Figure 2.3.

2.3.2.1 Heuristics

Quite often, individuals are unwilling to access and process all of the information necessary to make informed decisions on issues affecting privacy due to perceived or actual cognitive limitations, choosing rather to satisfy themselves with subpar solutions. Individuals constantly try to rationally maximize benefits but decision-making can only be rational within the limits of cognitive ability and available time (*Theory of Bounded Rationality*; Simon, 1982). Hence, the objectively irrational decision-making of the privacy paradox can be explained by an individual's cognitive limitations as they pertain to accessing and processing all relevant information which could well lead to a biased perception of privacy risks. A user's ability to access all the relevant information is bounded by nature, leading to a situation where the risks are deemed to be outweighed by benefits (Deuker, 2010). However, from their subjective point of view, the decision process may well appear rational (Flender & Müller, 2012). Even users with privacy concerns prove extremely reluctant to take the necessary actions to become informed, even when the information to protect one's privacy is made readily available (Acquisti & Grossklags, 2005). Information disclosure translates to a loss of control over that information and individuals find themselves in an information asymmetry which can be overcome through rational assessment. However, the factors that may play a role in that cognitive process are very difficult to aggregate, calculate and compare, requiring high cognitive involvement. As a result, the costs of adequate risk assessment can be perceived as unacceptably high leading individuals to rely on simple heuristics (Acquisti, 2004). In the context of social networking sites, teenagers in particular operate under bounded rationality and risk assessment takes place according to personal experiences pertaining to privacy invasion and not hypothetically in advance (Jia et al., 2015). Even if people theoretically have all of the necessary privacy-relevant information, they are unable to make proper sense of all of the information. This leads to the application of simplified mental models that often favor the benefits (Pöttsch, 2009).

Mental short-cuts allow individuals to come to decisions quickly while suppressing any urge to think about the next action (*Cognitive Heuristics*; Tversky & Kahneman, 1975). However, heuristics can lead to biases and an approach that has been successful in the past is no guarantee that it will prove suitable in another situation. Furthermore, heuristics hinder individuals from developing new ideas and alternative solutions. In their study of online privacy behaviors, Gambino et al. (2016) found a total of four positive heuristics (gatekeeping -, safety net -, bubble - and ephemerality heuristic) and four negative heuristics (fuzzy-boundary -, intrusiveness -, uncertainty - and mobility heuristic) that promote or inhibit information disclosure behaviors and possibly lead to the privacy paradox.

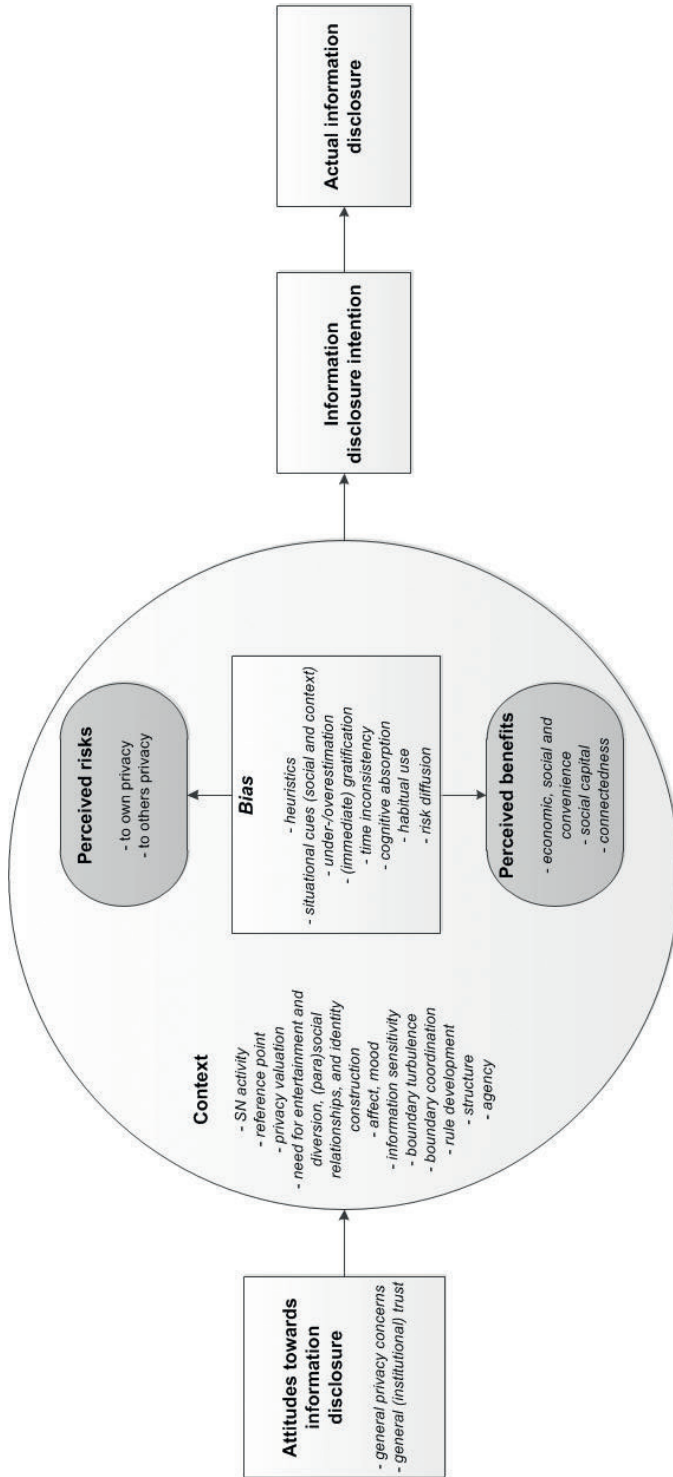


Figure 2.3 Overview of variables that play a role in the biased risk assessment within the risk-benefit calculation

In general, website cues that trigger affect heuristics (positive feelings about the website, for instance) have a direct effect on the trust and faith that eventually leads to online information disclosure (Wakefield, 2013). Also Sundar et al. (2013) found that privacy disclosure behavior in an online context is determined heuristically rather than systematically, which leads to a positive attitude toward information disclosure. In reference to cognitive heuristics and mental shortcuts, Kehr et al. (2015) concluded that the privacy paradox may result from misleading situational cues which bias cognitive valuation processes (e.g., affective thinking) and the prevalence of situation-specific considerations as compared to generic attitudes (*Extension to the Privacy Calculus Theory*). Eventually, privacy disclosure intention is determined by situational cues even if dispositional attitudes regarding privacy behavior are different to intention. Hence, privacy decisions are driven by situation-specific privacy assessment, general dispositions (i.e. general privacy concerns, institutional trust) and affect-based heuristics (quite often unconscious processes).

Cues-Filtered-Out Theory (Sproull & Kiesler, 1986, 1991) implies that individuals disclose more personal data in computer-mediated communication settings compared to face-to-face settings due to the absence of social and contextual cues leading to disclosure of information online despite general privacy concerns (Pötzsch et al., 2010).

When considering emotional factors during decision-making processes, individuals rely upon their feelings (mood, meta-cognition, emotion and body sensation), which generally leads to accurate responses but not always (*Feeling-as-Information Theory*; Schwarz, 1990, 2012). For instance, being in a good mood lets people evaluate targets or situations as more positive than they may be. Furthermore, judgments are based on feelings of ease or difficulty, as situation or targets that are easy to process are evaluated as more likely to be of less risk and more valuable. Being in a positive mood positively influences privacy protection attitude, whereas being in a negative mood increases risk perception (Kehr et al., 2014).

Considering interpersonal relationships, social life is more than individual acts, yet it is determined by social forces such as traditions, institutions and moral codes. Social structures determine human behavior but can also be altered as a result of perception, ignorance or replacement. Therefore, behavior is a balance between social structures and agency (the ability to act according to free will). The structure is achieved through a dynamic process as structure forms the basis for decision-making but is at the same time the outcome of it (*Structuration Theory*; Giddens, 1984). While making privacy decisions with regard to the usage of mobile applications, people are constrained by

structures (e.g., privacy requirements as a standard) which implies that they are unable to act entirely according to their free will (e.g., having general privacy concerns). This eventually leads to the negotiation of privacy by weighing costs against benefits. However, users are expected to accept certain requirements if they want to install and use a certain app. This eventually results in a contradiction between stated privacy attitudes and actual behavior to the extent that sharing personal information becomes perceived as normal in social life (Zafeiropoulou et al., 2013).

Likewise, decisions on what information to reveal and which to keep private are central to the *Communication Privacy Management Theory* (Petronio, 1991, 2002). Publication or retention is associated with particular risks and benefits. This process is guided by the subjective privacy boundaries that individuals have (as determined by an iterative process of rule development, boundary coordination and boundary turbulence), boundaries that are continually reshaped, depending on situation, context and communication partner(s). In this regard, rule development is assessed as a function of the nature of a given network (e.g., a network that is made of strong ties requires more information disclosure and therefore higher levels of privacy), whereby boundary coordination is tested by means of communication violation expectancy (e.g., if boundaries are violated, an individual might adjust their privacy settings in order to regain privacy). Hence, privacy decision-making is based on developed rules that depend on the perception of risks and benefits. However, this seems only be true for perceived privacy concerns and not for actual disclosure behavior, resulting in a “heat of the moment” paradoxical behavior (Sundar et al., 2013, p. 811).

2.3.2.2 Under- and/or overestimation of risks and benefits

Individuals tend to underestimate their own risk of privacy invasion while overestimating the chances that others will experience adverse events. This eventually leads to a belief that their own privacy is not at risk, a situation that can in turn eventually result in enhanced risk exposure (Acquisti, 2004; Flender & Müller, 2012). Furthermore, this lower risk perception might result in a laxer precautionary stance (*Optimistic Bias Theory*; Irwin, 1953). The tendency toward a reluctance to engage in privacy protection behavior against low probability but high impact/consequences events due to biased perception (event is less threatening than it actually is), underestimation of probability (as a consequence of little or no experience with the threat in question), unawareness of the threat or the costs of engagement are considered as too high is also discussed in the *Theory of Under Insurance* (Kunreuther, 1984).

The underestimation of future risks may lead to a tendency to underinsure oneself against these risks (Acquisti, 2004).

Likewise, according to *Third-Person Effect Theory* (Davison, 1983), individuals tend to overestimate the effect of media on others while underestimating the influence on themselves. As a result, individuals usually do not demonstrate the intended behavior as a response to the message. In this regard, the negative effects of information disclosure in social networks are mostly ascribed to others while they consider themselves the beneficiaries of positive effects only (e.g., building and maintaining relationships; Debatin et al., 2009). Although many users of social networks possess considerable privacy setting knowledge, they do not protect their private information properly as the perceived benefits outweigh any potential risks associated with information disclosure.

2.3.2.3 (Immediate) gratifications

In some cases, individuals encounter self-control problems as immediate gratification prompts atypical behavior which may be negative over the long term (*Immediate Gratifications*; O'Donoghue & Rabin, 2001; *Self-Control Bias*; Loewenstein, 1999). If there is choice, individuals usually choose a small benefit in the short term over a larger benefit in the longer term. If all choices are available over the long term, greater benefits will be chosen, even if these will occur later than smaller benefits (*Hyperbolic Discounting Theory*; Laibson, 1997). Individuals might have general privacy concerns; however this will not influence information disclosure behavior in the spur of the moment. Situational cues mitigate potential risks in the distant future and emphasize immediate benefits as users exhibit a tendency to favor immediate rewards on the short term at the expense of future risks due to a lack of self-discipline (e.g., using a search engine and getting a result immediately). This immediate gratification outweighs eventual privacy concerns resulting in poor risk protection by neglecting privacy protection technology even though they might encounter privacy violations in the future. Thus, individuals tend to heavily discount the low probability of high future risks (e.g., identity theft), resulting in a preference for almost instantaneous benefits (Acquisti, 2004; Acquisti & Grossklags, 2005; Deuker, 2010; Flender & Müller, 2012; Hughes-Roberts, 2013). During online transactions for example, users disclose private information in return for small benefits, even if their general privacy concerns are contrary to this behavior (Wilson & Valacich, 2012).

In their study on information disclosure on social network sites, Alashoor and Baskerville (2015) found that *cognitive absorption* (Agarwal & Karahanna, 2000; Agarwal et al., 1997) during social networking activity can overrule privacy-

related thinking as illustrated in the privacy calculus. The extensive use of social networking sites and the associated intrinsic rewards from that engagement can eventually lead to a flow state from being highly engrossed in such activities. This in turn can result in the kind of inappropriate behavior as it pertains to information disclosure that can lead to serious disadvantageous consequences affecting both career and private life.

Looking at the motivation perspective (*Uses and Gratification Theory*; Blumler & Katz, 1974; Katz et al., 1974), media use is actively executed in order to achieve and satisfy certain goals and needs along the dimensions of diversion and entertainment, building and maintaining relationships and identity construction. This assumes that individuals recognize their needs and how to satisfy them. Participation in an online social networks offers gratification among all three dimensions of goals and needs which eventually outweighs possibly privacy concerns, even when perceived privacy violations occur (Debatin et al., 2009).

2.3.2.4 *Difference between the judgments of risks and benefits*

According to *Prospect Theory* (Kahneman & Tversky, 1979), decision-making processes take place in two stages. During the editing stage, expected outcomes are ordered according to the basis of heuristics by setting a reference point. During the evaluation stage, outcomes below the reference point are considered losses and better outcomes as gains. However, individuals do not process information rationally. They value gains and losses differently as decisions are usually based on perceived gains rather than losses, with losses being judged more harshly than gains that might otherwise be judged equal. Hence, interaction with others via social networks (gain) can lead to privacy risk depreciation (loss) (Hughes-Roberts, 2013).

Making use of *Quantum Theory* (based on Busemeyer et al., 2006), Flender and Müller (2012) suggest that the choice between high and low privacy valuation and data disclosure and concealment are two incompatible types of preferences and an exchange of information between both cannot take place. Additionally, preferences are not predetermined but altered at the time an actual decision is made. Furthermore, privacy valuation and the associated privacy threats/risks are abstract concepts and data disclosure refers to concrete benefits. This explains why concrete benefits might often dominate abstract risks.

2.3.2.5 *Habit*

Repetitive behavioral patterns are addressed in the *Theory of Ritualized Media Use* (Rubin, 1984). (Social) media not only serves to satisfy information demands or entertainment needs, but can also be seen as a habit that is integrated into

daily routines. Such routines form a part of temporary structures and social rituals. Debatin et al. (2009) conclude that the use of social networking sites is ritualistic and strongly integrated into people's everyday lives by means of creating social capital and connectedness through a broad network (which would not be possible in an offline context) so that these benefits outweigh privacy concerns and prevent engagement in privacy protection behavior, even in cases of direct privacy violations. Quinn (2016) found that habit (habit was identified as one out of nine uses and gratifications) probably inhibits engagement in privacy management tools on social networks, despite increased experience with social networking, which eventually leads to a disconnection between privacy concerns and behaviors.

2.3.3 Little to no risk assessment

There are certain situations in which individuals possess little to no privacy-related information whatsoever, such as those where goal attainment nullifies all other considerations, or circumstances involving users who are unconcerned with privacy protection. Such situations result in significantly prevalent (perceived) benefits accompanied by negligible to no risk consideration. A complete risk-benefit calculation cannot take place under such circumstances and benefits are used as the sole reference point. Following, the different theories applied to the one-sided decision-making processes will be discussed. A comprehensive overview of all the variables that were mentioned in the following theories is presented in Figure 2.4.

2.3.3.1 *Value of desired goal outweighs risk assessment*

Privacy can be described as a dynamic process of regulating interaction with others. In this process, in-group and out-group considerations play a major role. Thus, based on internal states and external conditions, individuals determine their acceptable degree of openness. Privacy regulation should be enforced at an optimal level (an equilibrium between the desired level of privacy and the actual level). Here, trust plays an important role in the interaction regulation process which is defined by the self-imposed boundary (around a person) which in turn is determined by self-disclosure and a dyadic boundary (ensures the discloser's data security in the case of violation; *Privacy Regulation Theory*; Altman, 1975). According to Shklovski et al. (2014) repeated invasion of privacy boundaries can lead to a state of resignation (learned helplessness). People do accept privacy policies by app developers despite privacy concerns for the mere reason of having access to the app (mobile apps usually act according to an all-or-nothing principle that implies that total acceptance of the privacy policy

is inherent to use of the app) and because users are resigned to the fact that they possess little power to change the situation anyway. It would seem that the suppression of negative feelings as they pertain to the information sharing behaviors of some apps is simply part of the ‘cost of doing business’ – that is, the price one has to pay to use the app. An individual’s attitudes and behavior are influenced by others (especially by close or important friends). Individuals feel indirectly pressured to adapt their own behavior to achieve conformity with the admired peer group. Peer group pressure can result in either positive or negative reactions (*Conformity and Peer group pressure*; Crutchfield, 1955). In their study on online service providers, Flender and Müller (2012) found that peer group pressure negatively influences the privacy decision process. In order to be a group member, individuals neglect privacy concerns while disclosing information. Opting out is not considered an option because exclusion from the group is undesirable.

Looking at the interpersonal level, some forms of social collectives are determined by internalized emotional ties and implicit rules (*Gemeinschaft/Community*), whereas other collectives are determined by rational calculations and explicit rules (*Gesellschaft/Society*; *Theory of Gemeinschaft and Gesellschaft*; Tönnies, 2012). In social networks, people share private information because doing so is an implicit rule for belonging to a certain group. Although people are abstractly aware of data violation, these rational feelings cannot be translated into actual feelings of fear. As a result, the desire of belonging to a social network overrides any fears the consequences of data misuse might provoke (Lutz & Strathoff, 2011).

Furthermore, ‘impression management’ in social networks plays an important role. The *Extended two-component model of self-presentation online* (based on Leary & Kowalski, 1990) states that self-presentation is the process by which individuals try to control the impression that others form of them. This process is determined by two components: (1) impression motivation (the desire to create or re-create impressions in others’ minds; influenced by goal relevance, value of desired goal and discrepancy between desired and current self-image) and (2) impression construction (the process of creating this impression through change of behavior; influenced by self-concept, desired identity, role constraints, current or potential social image and target values). According to Krämer and Haferkamp (2011), the target value works differently in the online context compared to self-presentation in the offline context: either a wealth of information has to be given to broad audiences or vague information to avoid contradiction with others’ values. The former conflicts with privacy concerns and can therefore be seen as an inhibitor to online self-presentation.

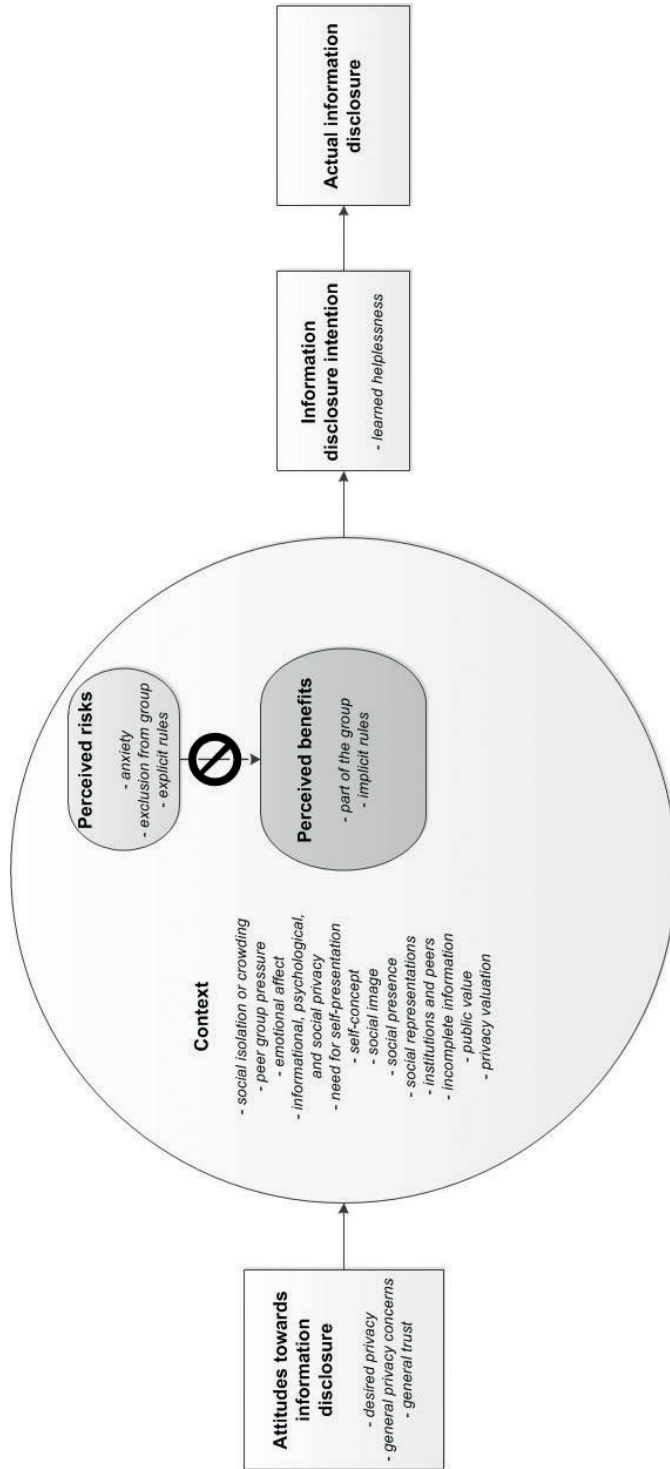


Figure 2.4 Overview of variables that play role in the decision-making process with no to little risk assessment

The latter places constraints on the goal to provide a detailed true self. Therefore, decision processes depend on the strength of impression motivation and privacy concerns that can eventually lead to a certain paradoxical behavior of self-disclosure.

2.3.3.2 *Privacy valuation failed*

Public Value Theory (Meynhardt, 2009) states that any organization contributes to society's wellbeing (objective facts), provided that individuals perceive their relationship to the public as either positively or negatively (objective facts are reflected in people's perceptions and subjective evaluation). If an organization is perceived as trustworthy regarding data protection but their public value is low, this organization does not contribute to the public value, unless general data protection is valued by the public. This partly explains the privacy paradox as people do not engage in protective behavior because they fail to value data protection (Lutz & Strathoff, 2011).

Relationships with others and how social roles, social support and social systems influence individuals' behaviors and the outcomes of such interactions are central to interpersonal relationships. Among others, to orient and master the social world, individuals form social representations and exchange such values, ideas or practices. As a result, new concepts are integrated into existing representations (making the unfamiliar familiar) by means of anchoring (the integration of new knowledge into existing representation) and objectification (making abstract concepts concrete by the creation of a new representation, e.g., the concept of privacy; *Social Representation Perspective*; Abric, 1996; Moscovici, 1984). According to Oetzel and Gonja (2011), the contradictory behavior of privacy protection occurs because privacy, as a concept, is not yet integrated into an individual's social representation.

2.3.3.3 *Knowledge deficiency due to incomplete information*

In game theory, one party is usually less informed than the other. In other words, not all parties know each other's values and rules. Users may be unaware of the importance of the data they disclose and what the consequences are of disclosing personal information (e.g., data are stored and processed by third parties). Due to this unawareness, eventual risks cannot be properly evaluated and the state of incomplete information prevents users from acting rationally and maximizing benefits (*Theory of Incomplete Information*; Harsanyi, 1967). This implies that users are lacking in privacy protection knowledge, at both the technological and legal levels; leading to misinterpretation of the likelihood of actual privacy violations and to inaccurate predictions of future hazards (Acquisti & Grossklags,

2005). Furthermore, not only are individuals ignorant as to the value of their data, they are unaware of the automated collection pathways and therefore unable to calculate the consequences of data disclosure at all. Consequently, costs are neglected and benefits preferred (Deuker, 2010; Flender & Müller, 2012). A viable evaluation of potential privacy threats requires processing quite a lot of information, information users often do not have and information that would likely prove superfluous anyway, as the probability of a future privacy violation is difficult for most to reasonably judge (Aquisti, 2004). In their study on app purchase behaviors, Buck, Horbel, Germelmann, et al. (2014) suggest that some costumers engage in subconscious purchase processes and a search for relevant information about issues such as data usage by third-parties does not even take place. The asymmetry of information possession and relying on mental short-cuts is strengthened even more by the operating systems in IOS (only one search result is prominently proposed to the users when searching for new apps) as a rule-of-thumb strategies play an important role in decision-making.

As previously discussed in the *Dual Process Model of Cognition* (Kahneman, 2003), decision-making is based on System I that is fast and automatic but vulnerable to influences that inhibit the rational decision-making process (produces intuitive concern, for instance), and System II that is rational and responsible for reasoning (produces considered concern, for instance). Relying on System I, individuals act on their intuitive concern without assessing the risks due to an incomplete understanding of it. Thus, no considered concern takes place and individuals are vulnerable to heuristic thinking (Phelan et al., 2016).

Considering the interpersonal level of decision-making while interacting with others, people share meaning and actions, and come to understand events in similar and certain ways (*Symbolic Interactionism*; Blumer, 1986). For instance, behavior on a social networking site can be described as a continuous process of information sharing and assessing reactions to this information from friends. By doing so, users become aware of the consequences of sharing information at a social privacy level but not at an institutional privacy level. Thus, users carefully monitor their self-presentation online but there is little or no interaction with the institutions that manage the information they disclose. This situation, in combination with low transparency as it pertains to data usage by provider companies, eventually leads to misinformation and misinterpretation of how third parties will utilize users' information and data disclosure despite privacy concerns (A. L. Young & Quan-Haase, 2013).

2.4 Discussion

The purpose of this chapter was to review prior research on the phenomenon of the so-called privacy paradox. Based on several theoretical approaches to decision-making, we were able to review the emergence of the privacy paradox through different lenses, differentiating decision-making according to a rational risk-benefit-calculation, a biased risk-benefit calculation and a decision-making process that involves no or only negligible risk consideration. We analyzed how these various theoretical approaches explain the apparent problem of paradoxical behavior (claiming to have privacy concerns but disclosing private information nonetheless) when engaging in online media activities, especially those pertaining to mobile computing (an important contributing segment when obtaining a more complete picture of the emergence of this phenomenon). In this final section, the main conclusions from the papers analyzed are drawn and implications for design and future research directions will be given.

2.4.1 Categories of decision-making in the context of information privacy

The systematic literature brought three decision-making categories effecting information privacy to light. First, decision-making can be divided into either a rational calculation of risk and benefits, or an irrational risk-benefit calculation characterized by biased risk assessment. Looking at the rational processes, individuals weigh costs against benefits, favoring gains over risks in most cases, such as using the service of an app or staying in contact via social network sites. Thus, information is given away in exchange for certain gratifications. Although users are aware there may be associated risks, compelling benefits or offers dull the perceived threats to privacy and safeguards are neglected. Looking at the irrational processes in decision-making, biases influencing the risk-benefit calculation play a role. Due to aspects such as heuristic thinking, (immediate) gratifications or time inconsistency, individuals are biased in their risk assessment, resulting in a distorted risk-benefit calculation, quite often tuned out to the advantages of associated benefits. The third category of decision-making describes processes in which negligible or no risk assessment takes place. Failed privacy valuations or information deficits for example, result in the risks associated with information disclosure being suppressed or even neglected altogether. All three categories of decision-making as it pertains to issues of information privacy might explain the discrepancy between stated attitudes and actual behavior, a phenomenon also known as the privacy paradox.

2.4.2 Rationality versus irrationality in decision-making

We believe, that to a greater extent, decision-making takes place on an irrational level rather than on a rational one, especially when it comes to mobile computing; suggesting that decisions are highly dependent on the context in which technology is used. The environment in which mobile applications are obtained and used means the decision-making process takes place much faster and on-the-go. This is partly supported by the studies which discuss the privacy paradox with regard to mobile applications (Buck, Horbel, Germelmann, et al., 2014; Deuker, 2010; Kehr et al., 2015; Zafeiropoulou et al., 2013) but challenging the assumption of a more rational view on the emergence of the privacy paradox (Y. Park et al., 2015; Pentina et al., 2016; Poikela et al., 2015; Quinn, 2016). However, we are in favor for a mixed approach when looking for potential solutions to overcome the privacy paradox as proposed by Lutz and Strathoff (2011) and Phelan et al. (2016). According to Phelan et al. (2016) design solutions should be adapted to different cognitive styles. The actual form such solutions might take remains ambiguous. To a large extent, it seems that individuals act on their intuition without assessing potential risks with regard to privacy and security intrusion, or they have considered concern but are constrained in their actions by external factors such as low transparency, user unfriendly design or consumer hostile privacy policies with all-or-nothing usage permissions. Sharing information (and using mobile applications) becomes a normal part of social life (Zafeiropoulou et al., 2013) and people are urged to accept certain requirements. The goal is to implement rational as well as irrational processes into design (backend and interface) so that decision-making eventually becomes self-determined. Only if this is fulfilled can a user's security and privacy concerns be diminished.

Consequently, we propose to raise knowledge and awareness by design, trigger heuristics through system and interface design (Gambino et al., 2016), support users through semi-automated or predominantly automated user-centered systems and user-friendly interface designs. Furthermore, we propose to make use of interactive software and social interaction to support the user as interactive software is typically perceived to be as trustworthy as a human communication partner (Berendt et al., 2005). Learning at the moment of experience (Jia et al., 2015) and empowerment of users in their self-determined decision-making should be a top priority.

2.4.3 Emergence of the privacy paradox is context-dependent

This review shows that during the last three years in particular, the issue of the privacy paradox and its emergence has moved into the focus of researchers, especially in conjunction with mobile application usage. However, the majority

of papers still focus on social networks and other online media. Comparing the results from social network studies with those focusing on mobile application usage, it seems that the privacy paradox within the mobile context is even more complex. This is possibly attributable to the routine, enhanced privacy policies, better technical support and comparatively long availability of online services such as social networks and e-commerce platforms, raising the question as to whether or not the same can be said of mobile applications. Restricting one's profile on social networks is the easiest way to protect against privacy threats and security intrusions. However, such protection measures are not easily accessible while downloading and installing apps, suggesting that the majority of users do not possess the expertise nor the experience to engage in what would be considered appropriate protective behavior. We would argue that the technical processes underlying mobile computing exceed the comprehension of most users. For this reason, the following question can be posed: Is the same generic term, 'privacy paradox' applicable to both stationary online activities and those considered mobile? A differentiation could open the door to a whole new area of interesting research possibilities.

2.4.4 Solutions to the paradoxical behavior of users

Concrete proposals designed to tackle the problem of paradoxical behavior—claiming to have privacy concerns but acting to the contrary—remain scarce. Current efforts are mainly focused on redefining guidelines for the process(es) that take place during the decision-making phase, such as the simplified possibility of restricting data access permissions during the installation of mobile applications. There are currently no viable solutions designed to span the gap between a user's intention and behavior. We believe that research into finding a solution to this problem deserves more attention. A movement to user-orientated design is needed in order to empower the user with the ability to make self-determined decisions on matters of privacy protection and online security. Shifting the reference point from 'not mine' to 'mine' goes along with higher risk perception which leads to the development of psychological ownership. This might elicit a higher valuation of private information, resulting in risk-averse decision-making (Kehr et al., 2014). Hence, individuals may be less vulnerable to disclosure influences due to their loss aversion and their sensitivity to loss (Baek, 2014).

In the context of context-aware mobile applications, Deuker (2010) propose privacy disclaimers at a situation-specific level to mitigate the effects of bounded rationality. Privacy awareness tools should empower users to make well-informed decisions with regard to their information disclosure (Pöttsch et al., 2010).

Furthermore, interface design should bring attention to such intentions in terms of mobilization (activating heuristics which protect the user; Kehr et al., 2014; Kehr et al., 2015). In their study on mobile websites, Zhang et al. (2014) concluded that a security warning with regard to the website resulted in increased threat perception as it pertains to private information, negative attitudes toward the service provider and lower tendency for future use. This puts an emphasis on positive user experience to promote privacy protection behavior (Kehr et al., 2014). In the case of privacy threats, psychological reactance behavior might be triggered (Brehm, 1966) as the individual tries to reach a certain state of restored autonomy through aversive reaction in a situation in which personal freedom is threatened (e.g., limiting or denying free choice by others). In other words, as a consequence of perceived privacy threats, users might try to regain control (freedom) by providing less personal data or even avoiding situations that could place them at potential risk (e.g., a cessation of app downloading). Creating privacy awareness in combination with tools that support users in their privacy decisions should help users to avoid paradoxical behavior (Deuker, 2010).

2.4.5 The special case of mobile computing

The use of mobile devices and the use of mobile applications in particular, falls within the realm of personal space extension: “...smartphones felt like intimate zones and extensions of their bodies” (Shklovski et al., 2014, p. 2347). This could explain why the majority of papers discuss the privacy paradox in mobile computing through theories at the intrapersonal level where cognitive, internal processes are used to process stimuli in order to behave a certain way. It seems that the downloading and use of mobile applications is—to a greater extent—self-referential in the sense that a user’s decision on whether or not to download a mobile application is done so in accordance with their personal preferences rather than those of their social group (with the exception of WhatsApp and other social networking apps not considered). This observation has not been researched to date, despite the fact that this might play an inherent role in solving the privacy paradox problem. Furthermore, individuals tend to organize their use of mobile applications according to goal orientation, with users showing a higher susceptibility to accepting privacy and security intrusions (e.g., a banking app) when compared to less important applications (e.g., a gaming app; Shklovski et al., 2014). However, research on this tendency is still scarce and should not be neglected when looking for solutions.

Furthermore, app stores for the Android operating system in particular, currently employ an all-or-nothing-policy, meaning that users have to accept all permissions in order to download a particular app; suggesting a deliberate,

systematic provocation of paradoxical behavior that could ultimately lead to overall acceptance of privacy risks as the stress that comes with a user's sense of inability and vulnerability dulls through continuous privacy invasion. Users express a desire for transparency and information control but are not able to act according to their needs for privacy (Shklovski et al., 2014). This makes it even harder to overcome the privacy paradox as the desire to own a particular app seems to outweigh potential risks all too often. This raises the question: To what extent is it possible to address this problem with the implementation of measures such as design prompts?

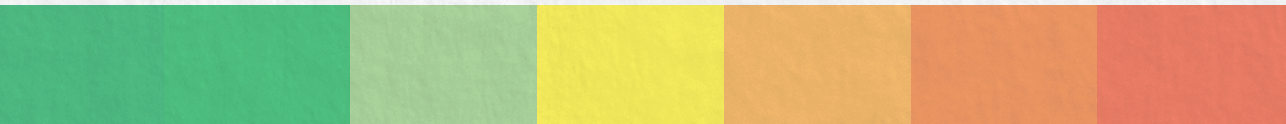
2.4.6 Research limitations

This review showed further limitations of the research to date. Attitude and intention are not actual behavior and actual behavior is seldom measured in the studies. Privacy concerns seem to be highly situation-dependent and can be described as a fluent concept that changes over time. However, most studies have researched privacy as a stable concept (Xu et al., 2010) or used conventional polls that are not suitable for studying online privacy (Baek, 2014). Dienlin and Trepte (2014) also rendered critique on prior measurements of the privacy paradox. When distinguishing between (i) information, (ii) social and (iii) psychological privacy, the results showed that the differentiation between these privacy dimension had a direct effect on the corresponding behavior (i: preference for disguising identity = less likely identifiable on a social networking site; ii: preference for restriction profile = more restriction applied; iii: preference for less personal information = less personalized profile). Accordingly, when distinguishing between privacy concerns and attitudes, applying the appropriate theory to the problem in question (here *TPB*) and differentiating on the above mentioned privacy dimension, the authors consider the privacy paradox as “a relic of the past” (Dienlin & Trepte, 2014, p. 295). However, we do not consider the privacy paradox as a relic of the past but we do believe that future research on the privacy paradox should try to measure actual behavior in order to get better insights into the problem.

2.5 Conclusion

The purpose of this study was to review prior research on the phenomenon of the privacy paradox. We highly question whether or not rational decision-making processes are the only suitable explanation for the discrepancies between privacy concerns, especially as it applies to mobile computing as decision-making in

a mobile environment is subject to environmental and circumstantial factors different from those encountered during desktop computing. When analyzing the design of mobile applications, we favor a mixed approach (rational and irrational decision-making) and design solutions should be adapted to different cognitive styles. Implementing cues into the design (backend and interface) is a necessary requirement for the empowerment of the user if data protection is to become more rational. However, attempts to theoretically explain and practically solve the problem of the privacy paradox are still scarce and we feel the subject deserves far more research attention.



3

Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources

Barth, S., De Jong, M. D. T., Junger, M., Hartel, P. H., & Roppelt, J. C. (2019). Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. *Telematics and Informatics*, 41, 55-69.
<https://doi.org/10.1016/j.tele.2019.03.003>

3.1 Introduction

At the time of publication, the number of smartphone users worldwide was just shy of 4.5 billion, with projections for the number of mobile phone users expected to reach the 5 billion mark by 2020 (Statista, 2016). Smartphone users store information and surf online, and by doing so collect (and distribute) large amounts of information. For billions of people around the world, the smartphone has become an indispensable daily companion. For many, the device remains within reach even while sleeping. Some researchers even argue that mobile phones can be seen as an extension to the human body (Shklovski et al., 2014). Because of their round the clock close proximity, smartphones can provide private behavioral information, including location, fitness, both on- and offline activities, social networking operations, and even audio-visual recordings (Aditya et al., 2014). However, most of this data aggregation is not voluntarily or consciously established by the end-user but initiated by business models based on data generation (Buck, Horbel, Kessler, et al., 2014).

By downloading and installing apps, smartphone users increase the risks associated with design flaws, malware attacks, and data theft. From a technical standpoint, the security and privacy risks associated with mobile applications have long been a cause for concern. By requesting irrelevant permissions, loosely defining permissions, or misusing permissions, combined with highly personalized data aggregation, mobile apps can and actually do provide third parties with sensitive data (Buck, Horbel, Kessler, et al., 2014; Egele et al., 2011; Enck et al., 2014). Most mobile users are unaware of these threats to their personal data or unable to understand the technical mechanisms behind data leakage (Acquisti et al., 2016). Consequently, information exchanged between electronic devices can be used for user monitoring, leading to a generally wary user attitude toward the credibility of a smartphone's privacy protection mechanisms. Although users might have ominous feelings when sharing information online, they still download and use apps in exchange for financial benefits, personalized services, or enjoyment in any way (Shklovski et al., 2014). Despite a foreboding feeling many users express, the number of mobile app downloads worldwide increases continuously (Statista, 2019). It seems that users download and install mobile applications without hesitation, even if apps require excessive permissions.

When it comes to privacy-related online behaviors, various researchers have drawn attention to the so-called 'privacy paradox', which refers to a discrepancy between users' attitude toward privacy and their actual behavior. Users claim to be seriously concerned about their privacy but in fact do very little to protect

their personal data (Acquisti, 2004; Barnes, 2006). Although the downloading and usage of mobile applications, for example, is often accompanied by a sense of insecurity and safety concerns, information sharing online is still on the rise (Zafeiropoulou et al., 2013). This seemingly paradoxical behavior can be explained by various psychological processes that take place during decision-making: (a) Users perform a risk-benefit calculation, guided by rationality, (b) they do show concerns but these concerns are overridden by factors such as desirability of the app, time constraints, or promised gratifications, or (c) they act on their intuition without assessing risk of information sharing online (Barth & De Jong, 2017).

This study aims to address deficiencies in the current privacy paradox literature. Firstly, when compared to the desktop online environment, research into the privacy paradox as it pertains to the mobile online environment is still very limited. Most available studies focus on social networking (Debatin et al., 2009; Dienlin & Trepte, 2014; Flender & Müller, 2012; Hu & Ma, 2010; Hughes-Roberts, 2013; Krämer & Haferkamp, 2011; Oetzel & Gonja, 2011; Poikela et al., 2015; Shklovski et al., 2014; Sundar et al., 2013; A. L. Young & Quan-Haase, 2013) and e-commerce activities (Acquisti, 2004; Acquisti & Grossklags, 2005; Motiwalla et al., 2014; Sundar et al., 2013; Wilson & Valacich, 2012), while research focusing on smartphone behavior and the use of mobile applications in particular remains scarce (Deuker, 2010; Oetzel & Gonja, 2011; Zafeiropoulou et al., 2013). Unlike when using traditional phones or computers, users are more prone to privacy intrusion in a (smart) mobile environment (Benenson et al., 2012; Williams et al., 2017), which underlines the need for more research in the mobile domain. Mobile application usage and the resulting data storage are continuously increasing. Considered alarming by many, the user is often excluded from decisions about which data can be shared and which should remain private. In order to support user empowerment, more research is needed into mobile application usage as it pertains to conscious, unintended or unwitting data distribution and sharing.

Secondly, actual behavior is seldom measured in studies addressing the privacy paradox. In order to gain better insights into the privacy paradox and offer an explanation of why people behave online as they do, we want to measure actual behavior instead of drawing conclusions based on stated intentions. This study aims to explore whether or not the privacy paradox is observable in actual behavior, making it more than a theoretical phenomenon which may be attributed to a measurement bias known as the intention-behavior gap.

Thirdly, research has shown that a knowledge and awareness gap can lead to a certain paradoxical behavior as it pertains to information disclosure online.

For most users, technical processes that run in the background when doing business online are neither visible nor understandable. Consequently, technical skills (e.g., downloading an app) cannot be equated with technical literacy (e.g., understanding the data flow processes in play while downloading an app), leading to a situation in which users make use of online services despite concerns about privacy or security issues (Liccardi et al., 2014). In order to mitigate the potential influence a lack of technical know-how might have, we studied the privacy paradox among users with a high level of technical expertise and awareness regarding online privacy and security.

Fourthly, financial restrictions are considered a significant factor in this paradoxical behavior as well, especially as they pertain to mobile computing and more specifically, when installing apps on smartphones. Users have a tendency to not buy their apps, even if they cost mere cents (Liccardi et al., 2014). App developers often use advertising or re-use app data for other purposes to generate revenues. It is a proven fact that free versions of many types of apps require a broader scope of permissions—often unrelated to the apps’ functionality—than purchasable versions of similar apps (Chia et al., 2012), opening the door to user data misuse. In order to compensate for the possible effects of financial restrictions, participants in our study were provided with a certain amount of money that could be used for, among other things, an app purchase.

Consequently, this research aims to explore users’ actual behavior when installing an app on their smartphone, compensating for any influences attributable to technical knowledge and privacy awareness deficits while mitigating the influences of financial restrictions, leading to the following research question: *To what extent do technically skilled mobile phone users, in a setting that is controlled for a prominent role of financial considerations, show a discrepancy between perceived privacy concerns and actual privacy-related behavior while downloading and installing a mobile app?*

Below, a review is given of the relevant literature on the current situation of data handling and privacy threats, the privacy paradox and privacy concerns as they pertain to the use of mobile computing technology. This is followed by a description of, respectively, the design of our study and the results found. The chapter concludes with a discussion of the main findings and their implications, a reflection on the limitations of this study, and general conclusions.

3.2 Theoretical framework

The purpose of this study is to further investigate some of the factors that play a role in the decision-making process of consumers while downloading and installing an app. The main focus is on the privacy and security precautions users might take during the installation and usage of an app. Many definitions of privacy exist and the concept of privacy has changed over time and with continuously evolving new (smart) technologies. When we talk about privacy, we refer to information privacy and more specifically to online privacy. It involves deciding what personal information may be revealed to others and understanding how this personal information is obtained by others and how other parties make use of this information (Westin, 2003).

3.2.1 Mobile users' privacy and security behavior

The new age of information technology, especially with regard to mobile computing and the associated collection of huge amounts of private data, underlies a substantial business model as organizations might profit from extensive data gathering by trading personal data with other parties, developing new markets and services (Acquisti et al., 2016). Users seemingly disclose information and share their data without hesitation. Relatively vague legislation enables organizations to trade consumer data in order to reduce costs, enhance returns via advertising, and offer personalized services. The legality of the data handling practices conducted by many organizations is considered a grey zone (Spiekermann et al., 2015). Through sharing data publicly on the internet, the line between legitimacy and invasion of privacy is blurred by the users themselves. This does not mean that users are satisfied with the current situation. On the contrary, generally speaking, most users are only pro data sharing if they are consciously involved in the data exchange process, or if the extent of their personal data processing is considered acceptable (Spiekermann et al., 2015). Nevertheless, one can observe that those users who are ill-informed about data handling in particular, are more prone to sharing information. Acquisti et al. (2015) argue that users have to deal with fuzzy boundaries regarding online interactions. From a user perspective, the actual goings on in the cyberworld are unclear. Consequently, the privacy experience and the effects of privacy intrusion are not felt directly. If people's personal space is violated in the offline world, many feel immediately uncomfortable with the situation.

However, if the proximity related to privacy preferences is breached online, the effect of this violation is less tangible. Benenson et al. (2012) found that users do not translate their knowledge about desktop devices to the mobile context,

even though security and privacy threats are similar. Protection mechanisms (e.g., the installation of a firewall) are prominent on computers but not in mobile environments. Furthermore, the technical processes running in the background of certain apps cannot be ascertained via the permissions; even if users take the time to read and understand the agreement covering such permissions. Liccardi et al. (2014) illustrated this problem using the example of a weather app that asks for access to the internet and location information in order to deliver accurate information on current weather conditions. In such cases, users are often unaware that the information in question might also be used for other purposes such as tailoring advertising from third parties. Williams et al. (2017) found that users perceive IoT devices as less privacy-respecting compared to non-IoT devices, possibly due to hidden data collection or unknown technical processes underlying such smart devices. Privacy valuation seems to be lower in a mobile environment and users seem to value perceived benefits above perceived risks. Although Chin et al. (2012) found that users are less likely to share sensitive data (e.g., health data) on their mobile phones than on their laptops, they use other sensitive services (e.g., location based) because of the perceived benefits of such services. However, the data obtained in a mobile environment are much richer than those in a desktop environment: Portable devices permit service providers to grasp not only few glimpses of users' daily lives but to get a fine-grained picture about daily activities and even inner thoughts and feelings (T. Wang et al., 2016).

3.2.2 The privacy paradox

Research into online privacy shows that users are interested in privacy protection but that their privacy concerns rarely translate into actual behavior (Barth & De Jong, 2017; Joinson et al., 2010; Tsai et al., 2006). The discrepancy between expressed privacy concerns and actual, contradictory behavior is known as the privacy paradox: Users claim to have privacy concerns but do not behave accordingly as they engage in risky downloads and seemingly reveal private information without hesitation.

When examining factors influencing this contradictory online behavior, different explanations emerge, many of them focusing on general internet activities such as e-commerce or social networking. According to Kokolakis (2017), users might show distinct privacy behavior in different contexts, suggesting that privacy behavior is highly context-dependent. Bergström (2015) found that privacy concerns are increasingly present the more personal an application is, but nevertheless users tend to negate their privacy concerns for the mere reason of enjoying certain services, including mobile computing. In their study on general online behavior, Hoffmann et al. (2016) considered privacy

cynicism as an explanation for the paradoxical behavior users show online, even intensifying online self-disclosure. Few studies discuss the privacy paradox in a mobile computing context, although the privacy paradox seems to be even more complex for mobile app usage as users might be unable to employ the same protection mechanisms they would apply in a desktop environment. In their study on smartphone security, Volkamer et al. (2015) concluded that lacks of awareness, concern, self-efficacy and compulsion prevent users from adopting smartphone security apps. As far as mobile app users are concerned, restricting one's profile as one can on social networks, or adjusting privacy settings is not an option. In this line, Pentina et al. (2016) found that future app use is not limited by privacy concerns. Here, social and informational benefits seem to be the driving forces behind mobile app adoption. In their study on the privacy-personalization paradox among Chinese participants, Guo et al. (2012) found that trust highly mediates the effect of privacy concerns and personalization concerns on intention to adopt a mobile health service, suggesting that benefits of a health service should outweigh privacy concerns. Morosan and DeFranco (2015) found that positive emotions influence the perception of absolute value of mobile hotel apps, obscuring the sight of privacy concerns and enhancing the willingness to disclose personal information. However, ensuring secure data handling methods should be the top priority in order to overcome fear of data misuse and eventually paradoxical behavior. In this line, Sutanto et al. (2013) concluded that a privacy-safe design enhances app usage (process gratification) and interaction with the app (content gratification) and at the same time lowers feelings of personal data being breached, eventually diminishing the privacy paradox.

Furthermore, a lack of technical expertise regarding the conditions and procedures behind mobile computing, the seemingly impenetrable app market business models, and the economics of privacy may play an important role in strengthening the mechanism of the privacy paradox. In their study on users' perceptions of privacy in the IoT domain, Williams et al. (2017) concluded that price and functionality outweigh privacy, leading to the adoption of IoT devices or services despite having privacy concerns. In general, their study showed that users of smart devices engage less in privacy protection tools than non-IoT users. Several underlying psychological processes may eventually lead to people's contradictory behavior: (1) decision-making is based on a rational weighing of benefits and risks of downloading and using apps, (2) the weighing of benefits and risks is biased by psychological processes through extenuating circumstances such as immediate gratification, time constraints, or information deficits or overload, and (3) the risks involved in downloading and using apps is not even part of people's considerations (Barth & De Jong, 2017).

All three processes may lead to a situation in which benefits overshadow risks. Threats to privacy, such as the excessive collection of data and the installation of malware causing data leakage, monetary loss or disclosure of identifiable data to unauthorized parties are accepted for nothing more than the satisfaction of mobile computing. Users of Android operating systems are often confronted with an all-or-nothing choice in which they must accept all of the requested permissions if they want to download a specific app to their smartphone. Consequently, though users may have permission-related concerns, the desire to use an app seems to outweigh any risks associated with the installation process. Furthermore, searching for and actually downloading and installing an app is characterized by habitual, impulsive or limited processes due to the technical environment in which app purchases take place. The design of app stores and the way apps are presented do not seem to call for high involvement of users; apps are purchased rather automatically, without sufficient attention to privacy considerations (Buck, Horbel, Kessler, et al., 2014).

3.3 Method

3.3.1 Research design

Gaining in-depth insight into the decision-making process of mobile users requires intensive collaboration with actual users. In order to accomplish this, a manageable user group with substantial technical knowledge and skills was studied intensively. For this study we recruited Computer Science students who participated in a Master course on Cyber Crime Science in 2017 and 2018. We placed them in an experimental setting, in which they received sufficient money for buying a mobile app within a certain category and had to choose from five alternative apps, with varying degrees of privacy threats. Instead of merely expressing behavioral intentions, our participants were obliged to download the app on their mobile phone and use it for a week, so that they could write a review of the app. Apart from the actual app selection and downloading, we also administered questionnaires focusing on technical knowledge and skills, privacy awareness, and download considerations, and analyzed the app reviews the participants wrote. The entire research covered a period of three weeks. Figure 3.1 gives an overview of the overall design of our study.

The study was approved by the ethical committee of EEMCS faculty, University of Twente. All personal identifiable information was anonymized so that data from participants in all three parts of the study could be connected.

We introduced the research to the participants as a user experience experiment focusing on mobile apps.

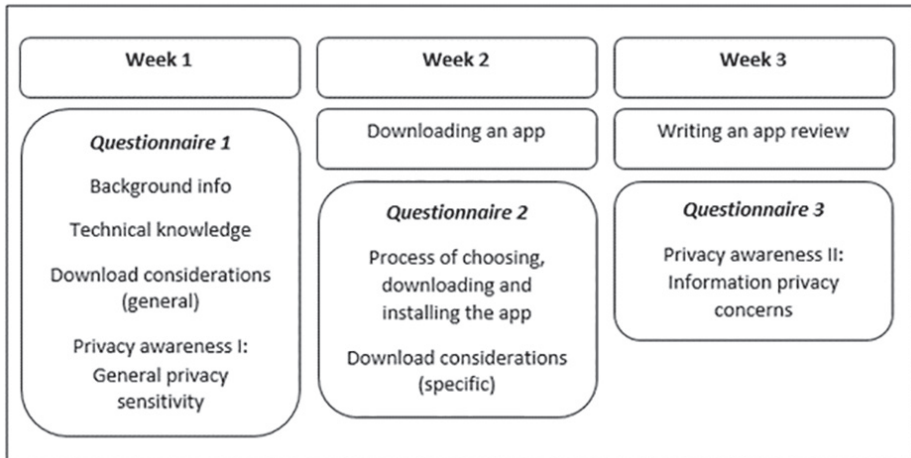


Figure 3.1 Overview of the research

3.3.2 Stimulus materials: Selection of apps

For our study, five apps out of two categories were chosen: utilitarian versus hedonic apps. Utilitarian apps serve to fulfill users' task completion needs, while hedonic apps serve to fulfill entertainment needs (Hazarika et al., 2016). We chose apps from these two categories to explore whether users' privacy considerations differed between them. Hazarika et al. (2016) expect differences in consumer addiction and frustration between utilitarian and hedonic apps. In a study on mobile computing, Wakefield and Whitten (2006) suggest that perceived enjoyment (attributable to hedonic value) enhances cognitive absorption, which in turn promotes usage behavior of a mobile device. Based on these findings we assumed that using an app for a utilitarian or hedonic purpose might lead to differences in privacy evaluation. Entertainment needs might make users less cautious when it comes to privacy valuation, as fun and enjoyment might be valued as more important than privacy protection.

As a typical representative in the utilitarian category we chose a to-do-list app; for the hedonic category we chose a tower defense gaming app. The selection of app categories was based on the study of Heinonen and Pura (2006) and a pre-test with 25 participants. Only apps that were actually available in the Google Play Store (Android Version 7.1.1) were selected.

Within both categories, we selected apps with similar core features and an increasing number of permissions requested. Hence, all apps from the to-do-

list category had to (1) use the term ‘to-do-list’ in the title or in the description of the app, (2) provide creation, management and categorization of tasks, (3) be able to set reminders, and (4) share notes and (5) have a user rating of four or higher out of five stars. For the gaming app, all five apps had to (1) belong to the tower defense game category, (2) provide different gaming levels to be achieved, (3) protect someone or something from someone or something, and (4) have a user rating of at least four or higher out of five stars. The apps in both categories had to differ in terms of permissions requested, ranging from a purchasable app that asked zero permissions to free of charge apps that asked up to seven permissions (identity, contacts, phone, photos/media/files, WIFI connection information, device ID, and call information). The increasing number of permissions had to be similar between the two categories. Another selection criteria for the apps was that the permissions requested had to be shown to users before they actually install the app on their smartphone. The apps were selected in collaboration with a computer science expert and a privacy intrusiveness score was assigned to them, ranging from not intrusive (zero permissions requested) to very intrusive (between five and seven permissions requested). The privacy intrusiveness scores were used to classify the apps according to the number of permission requested: The more permissions requested that were unrelated to the apps’ core functionality, the higher the intrusiveness score. An overview of the selected apps, requested permissions, and privacy intrusiveness scores can be found in Appendix 3.1.

3.3.3 Procedure

The first part of the study took place during the first lecture of the Master course. After reading and signing a consent form, participants started by filling out the first questionnaire, focusing on participants’ background information, technical smartphone knowledge and skills, privacy awareness, and download considerations in general. The privacy awareness questions in the first questionnaire were rather general, so that participants were not prompted that privacy was the main point in the research.

In week 2, the experimental part of the study took place. We explained to the participants that we would like them to evaluate the user experience of a specific app and therefore asked them to download an app in a specific category, use it, and write a short review. Participants were randomly assigned to one of two categories of apps (to-do-list vs game). Before starting with the selection of an app, participants were given 10 euros with the instruction that they could use the money for an app purchase but were also allowed to keep it. We explicitly informed participants that they were totally free to decide which app to download

and install on their smartphone. We requested all participants to use the app over the course of the following seven days to be in the position to write a comprehensive review about the app. After having installed the app, participants were asked to describe the process of choosing, downloading and installing the app and to explain why they had chosen for this particular app (and not for the other options). After that, participants filled out the same questionnaire about their download considerations, this time in relation to the downloading and installing of the specific app.

In week 3, participants were asked to write a review of the app they had downloaded and used over the course of the previous week. This was deliberate, as writing a review significantly differs from describing individual decision-making processes when downloading an app, or justifying one's selection for a particular app. User reviews can be defined as an implicit form of user-centered communication focusing on 'perceived ease of use' or 'perceived usefulness' (F. D. Davis, 1989; Vasa et al., 2012). Our main goal of the review assignment was to verify if privacy related factors played a role in the creation of co-value for other users (Tan & Vasa, 2011). We also included three control questions to ensure that subjects actually downloaded the app. After writing the review, the participants filled out a third questionnaire about their privacy awareness.

3.3.4 Measurement instruments

In questionnaire 1, participants answered several *background questions*, regarding demographics (age, gender, study program) and general app usage. One item was adapted from Yang (2013), investigating the number of apps someone has ever installed by his/herself on their smartphone (open question). Furthermore, we asked the participants which categories of apps they used on their smartphone: Participants could choose from 36 categories (as mentioned in the Google Play Store) in total.

To verify participants' *technical knowledge and skills*, four self-report questions about their knowledge and skills, based on Kraus et al. (2014), were asked in questionnaire 1. For instance, participants had to indicate whether they are able to protect themselves against data misuse. This scale's Cronbach's alpha was .61. In addition, questions were asked in questionnaire 1 about the extent to which they showed privacy and security related behaviors on their mobile phone. The questions were adapted from Androulidakis and Kandus (2011), and focused on behaviors such as using Bluetooth functions and running an antivirus app on their smartphone.

In questionnaire 1, the Westin Privacy Index, based on Kumaraguru and Cranor's (2005) survey of Westin's studies, was used to obtain a general

privacy sensitivity index of the participants and to investigate *general privacy awareness*. Three items concerning perceived loss of control over private data, data handling practices by third parties and laws and regulations pertaining data protection were measured on four-point Likert scales (1 = strongly disagree; 4 = strongly agree). The calculation of the privacy index was done according to Kumaraguru and Cranor's guidelines (2005).

In questionnaire 3, participants filled out a Mobile Users' Information Privacy Concerns Scale, which was adapted from Xu et al. (2012). The scale consisted of nine items measured on seven-point Likert scales (1 = completely disagree; 7 = completely agree). The scale contained the following constructs: perceived surveillance ($\alpha = .91$), perceived intrusion ($\alpha = .84$), and secondary use of personal information ($\alpha = .80$). To enhance Cronbach's Alpha value of the construct perceived surveillance, one item was deleted, ending up with eight items for this scale.

To investigate participants' *download considerations*, we presented them with a list of 18 aspects that could be considered when deciding to download an app, both in questionnaire 1 (in general) and in questionnaire 2 (after downloading the specific app). By doing so, we explored if the perceived importance of factors differs when participants give general estimations versus when they have just made a decision to download an app. The questionnaire contained seven privacy and security related factors (e.g., trustworthiness of the app, and number of permissions requested), and 11 other factors (e.g., prior experience with the app, price, and design). Participants answered on seven-point scales (from 'almost never true' to 'almost always true'). The complete set of questionnaires can be found in Appendix 3.2.

3.3.5 Participants

All participants ($N = 66$) were university students with a technical background in the Netherlands: 60 students followed the Computer Science Master program, three students attended the Electrical Engineering Master curriculum, two students studied Technology Management, and one student was completing a Business Information Technology Master. Their ages ranged between 19 and 55 years old, with an average age of 23.5 years ($SD = 4.6$). At 83% male and 17% female, the gender distribution in this study reflects the current male-female ratio in the technical Master programs. Obtaining a viable participation number large enough to compensate for any gender issues was deemed unfeasible. For the analysis of the experimental part of the study, only data from participants running an Android operating system (77%) on their smartphones who had filled out all three questionnaires were included in the analysis ($N = 39$).

Since owning their devices, participants reported having installed an average of 53 apps (SD = 41) on their smartphone. On average, participants had already owned a smartphone for 7.1 years (SD = 2.8). They spent an average of 2.6 hours a day using their mobile devices and interacted with an average of 10 mobile apps (SD = 6.1) during a normal week. In general, participants considered themselves to be experienced in using mobile apps (M = 6.0, SD = 0.7; measured on a seven-point Likert scale).

Apps with a communication purpose (e.g., 'WhatsApp') were the most popular app category participants use on their smartphones (95%), followed by 'business' (77%; e.g., Office and PDF) and 'maps and navigation' (76%; e.g., GPS navigation). In contrast, categories such as 'medical' (8%; e.g., 'DocCheck') or 'house and home' (8%) were rarely used.

Usually, when looking for new applications, 52% consulted an app store in general, whereby 46% indicated Google Play Store to be the main app store used to gather new info on new apps on the market. Furthermore, 6% of participants based their decisions for downloading new apps on reviews, 14% used search engines to gather info on new apps available and 2% relied on their own experience. Of the participants, 98% indicated that their smartphones had never been lost or stolen. Only 3% of the participants claimed that they lent their smartphone to others without hesitation, whereas 64% would only do so for a short while physically present; almost one third of the sample claimed that they never shared their smartphones with others (33%).

3.3.6 Analysis

Quantitative data analysis was divided into two parts: For the first questionnaire, the analysis was done on the basis of the 66 participants that completed the questionnaire. However, for the analysis of the downloading process and the second questionnaire, subjects that did not run Android operating system on their smartphones and did not go to the Google Play Store to obtain the app, were removed from further analysis. This resulted in a sample of $N = 39$ participants that took part in all three parts of the study.

The description of the process of choosing, downloading and installing the app and the reviews were analyzed qualitatively by means of a coding scheme, aiming at identifying concepts that relate to the question of interest (here: willingness to download an app). The coding scheme was based on the 17 download considerations asked for in the first and second questionnaire. Based on an open coding process, which means identifying other key factors that are not pre-determined or based on existing literature and concepts but emerging from the textual data directly, 13 factors were eventually added to the coding scheme

(e.g., color, interface, name of the app). To ensure inter-rater reliability and for validation of the coding scheme, the two open questions from the second part of the study and the review were coded by two independent coders. Subsequently, the coding of the two raters were compared and an agreement score, called Cohen's kappa, was calculated. Cohen's kappa measures the agreement between raters in their codings, correcting for chance agreement. The Cohen's kappa was .88, which indicates substantial agreement between the two coders.

3.4 Results

Below, the results of the quantitative and qualitative analysis of the data will be presented. In the first two sub sections, overviews will be given of, respectively, participants' technical knowledge and skills and their privacy awareness. After that, the focus will be on their app downloading and installing behavior. The final sub sections will, respectively, focus on the results regarding download considerations and the reviews.

3.4.1 Technical knowledge and skills

The following results give an indication of the participants' level of technical knowledge as it pertains to smartphones and how participants actually deal with situations that might affect smartphone privacy and security. Overall, the results show that participants confirmed to have a relatively high knowledge level regarding the technical aspects of their smartphones (see Table 3.1). This reflects the technical backgrounds of the participants. Looking at the cumulative percentages of disagreement or agreement, participants considered themselves to be highly knowledgeable and well-informed.

Asking participants about a smartphone's technical specifications, 73% indicated to know where you can find the mobile phone's IMEI (International Mobile station Equipment Identity: a 15-digit code that uniquely identifies valid smartphone devices, comparable with a PC's MAC address). In case of loss or theft, the IMEI code can prove helpful in smartphone retrieval. Thus, for security reasons, it is helpful to know its IMEI. After (re)starting their smartphones, 92% made use of a PIN code for unlocking their SIM card and almost the same number of participants used a PIN code or password to unlock the phone itself/screen-saver (85%). Locking one's SIM card with a PIN code supports data security as this inhibits unauthorized access. Furthermore, entering the wrong PIN code several times may lock the SIM card permanently. Activating the locking-function of the screen-saver is an additional security protection mechanism.

Table 3.1 Participants' knowledge about mobile phone's privacy and security (N = 66)

	<i>M'</i>	<i>SD</i>	Disagreement (%)	Agreement (%)	Undecided (%)
Communication through mobile phones is safe.	3.4	1.7	61	38	1
I am aware about how technical characteristics affect security.	5.1	1.5	17	76	7
I know how to protect myself against data misuse.	5.2	1.6	15	82	3
I know how to protect myself from malicious apps.	4.9	1.5	23	73	4

Note. 'Measured on seven-point Likert scales (1 = completely disagree, 7 = completely agree).

From the sample, 38% never used the Bluetooth function of their smartphone, whereas 52% claimed they do use Bluetooth but only for a specific purpose, after which the Bluetooth function is deactivated. Only 9% had their Bluetooth permanently activated and thereof 8% claimed to use to have their device visibility setting set to 'not visible'. The current security processes employed by Bluetooth wireless technologies are considered adequate, but the protocol might present a door-opener for attackers, such as worms and viruses or 'bluebugging'.

Relatively few participants employed an antivirus app on their mobile device (27%) and even fewer ran static analysis apps to monitor malicious code patterns, to inspect control flow between apps, or to review requested permissions (17%). While these results may appear low, it should be noted that, compared to the results of Androulidakis and Kandus' (2011) study, the antivirus numbers in this sample are approximately 100% higher.

When analyzing any sensitive data owners store on their smartphones, 5% stored passwords without encryption, 32% stored passwords such as credit card passwords with encryption, and 62% indicated to never store passwords on their mobile devices. The protection of personal data through encryption technologies reduces the risk of surveillance by third parties and enhances both anonymity and privacy. However, 83% of the sample confirmed that they have sensitive personal data such as photos, videos or audio recordings on their smartphones.

Of the sample, 70% claimed to create backup copies of their phone's data; 3% of them indicated that such backups follow no particular schedule (e.g., after resetting the mobile phone or a cleaning action), 20% backed up their data continually, for instance in a cloud, 12% ran a backup on a daily basis, 6% weekly, 12% monthly, 5% quarterly, 8% semiannually, and 5% annually. Backup copies are considered essential by security experts as security violations can often result in data losses.

3.4.2 Privacy and security awareness

A calculation of privacy sensitivity scores based on Westins' Privacy Index (H. Taylor, 2003) showed that privacy was not an issue for 9% of the sample. These individuals showed little or no hesitation when it comes to disclosing private information as they considered the benefits of information disclosure to far outweigh any potential risks. Of the sample, 41% can be assigned to the group of pragmatists. They weighed the potential costs and benefits of information disclosure before progressing. Pragmatists are willing to disclose information if they perceive their privacy protection expectations met and the party in question to be trustworthy. The majority of the participants can be described as fundamentalists (50%). Fundamentalists are at the maximum end of privacy concerns. They put responsibility for privacy protection at the individuals' level and require proactive refusal of information disclosure by users.

With regard to user privacy and security issues, the results showed that the significant concerns participants have in general are intensified when it comes to their smartphone usage. Generally speaking, participants were concerned that mobile apps monitor their activities and that often too much personal information is collected. The level of privacy intrusion by mobile apps in general was perceived to be relatively high. The level of concern regarding secondary use of personal information was even higher. A comprehensive overview of the results is given in Table 3.2.

Table 3.2 Overview of mobile users' information privacy concerns (MUIPC; Xu et al., 2012)

Construct	Explanation	M	SD
Perceived surveillance	Practice of data collection, track and profile mobile users	5.5	1.3
Perceived intrusion	Violation of physical and informational space	4.8	1.4
Secondary use of personal information	Unauthorized data usage for secondary purpose	5.3	1.1

Note. Measured on seven-point Likert scales (1 = completely disagree, 7 = completely agree).

3.4.3 Actual privacy-related behavior

Below, the results from the experimental part will be presented. Although there were less intrusive alternatives available, many participants selected the most intrusive mobile app, i.e. the app that requested most permissions that did not relate to functionality (28%). Furthermore, 49% downloaded the app that was analyzed as intrusive, and another 18% chose the app that was somewhat intrusive. Only 5% of the participants decided to buy an app that did not ask for

any permissions. One participant withdrew from downloading and installing an app due to privacy concerns. Table 3.3 provides an overview for both types of app. We can observe that for both app categories most participants chose the intrusive or very intrusive app, despite having the extra money for buying a non-intrusive app. The differences between the type of app (utilitarian vs hedonic) were not statistically significant.

Table 3.3 Relation between app intrusiveness and participants' download decisions for the two types of apps (in percentages) ($N = 39$)

Type of app	Number of requested permissions				
	0	1-2	2-4	4-6	5-7
To-do-list	0	0	18	55	27
Game	12	0	18	41	29
Overall	5	0	18	49	28

Note. See Appendix 3.1 for a more detailed overview of the categorization of app intrusiveness.

3.4.4 Participants' download considerations

Table 3.4 provides an overview of participants' download considerations in general (before the assignment to download a specific app) and immediately after downloading the app. Based on these results, several observations can be made. The first is that in both questionnaires the privacy and security related considerations did not play a prominent role. In the first questionnaire, only trust in the app and the relation between permissions and app functionality were in the upper half of the considerations. In the second questionnaire, only the number of permissions requested and the relation between permissions and app functionality were in the upper half.

A second observation is that specific and relatively easily judgeable considerations were more prominent in the second questionnaire than in the first. General considerations participants had in the first questionnaire when reflecting on downloading apps, all requiring the combination and weighing of several features, were often replaced by much more straightforward considerations in the second questionnaire. The top three considerations in the first questionnaire (functionality, usefulness, and trust) were replaced by price, ratings, and design in the second questionnaire. Especially the role of design is remarkable: In the first questionnaire it took a 14th rank. Within the privacy and security domain, trust in the app went down from a third rank to the 12th position, and some of the more specific considerations, most notably the number of permissions requested got higher ranks. This tendency seems to reflect that it may have been hard for participants, despite their technical knowledge and skills and their privacy

awareness, to incorporate privacy and security considerations in their download decisions.

Table 3.4 Mean ranks of download considerations before and after downloading and installing the app

Considerations	Before installation	After installation
Perceived functionality of the app	1	4
Perceived usefulness of the app	2	8
* Trust in the app	3	12
Price of the app	4	1
Ratings given by others (“star system” in the app store)	5	2
Reviews about the app (written by others)	6	6
* If permissions relate to functionality	7	7
Number of downloads (as indicated in the app store)	8	10
Recommendation of others	9	15
Prior experience with app	10	17
* Number of permissions requested	11	5
* Clarity of permissions (if permissions are understandable)	12	9
Familiarity with the app	13	16
Design of the app	14	3
* Privacy conditions (e.g., information disclosure to third parties)	15	13
* Readability of permissions (comprehensibility for the user)	16	11
* Security conditions (e.g., if data protection is ensured)	17	14

Note. Privacy and security related factors indicated with asterisks

A third observation involves a mismatch between the considerations mentioned by the participants and their actual downloading behavior. The number of permissions requested ranked 5 in the second questionnaire. However, many participants ended up downloading the app that asked for the most permissions. Thus, it appears that participants claimed to have privacy concerns that are not reflected by their actual behavior. Privacy considerations are anchored in users’ mindset but do not manifest themselves as a top-priority.

When describing their decision-making process while choosing for an app, downloading the app and eventually installing it, about 8% of the participants mentioned the factor *permissions*. Of these, the majority indicated that permissions asked by the app played a role in their decision whether or not to download it as they “looked at [...] the permissions it required,” “checked permissions,” or “secondly I looked at the permissions the app needed.” However, participants who claimed to have looked at permissions asked by the app often chose to download the app that was judged as somewhat intrusive rather than the paid app that asked for zero permissions nonetheless. Some

participants appeared to be discouraged by the required permissions and withdrew from downloading an app even though they thought it to be the most attractive one *“because it required a lot of privacy sensitive information.”* Some comments were more nuanced and specifically mentioned if permissions relate to functionality by reviewing the privileges the app asked for *“to check whether they were logical for this kind of app.”* Sometimes, users felt uncomfortable because of the requested permissions as one participant stated that he had *“inspected the permissions and was baffled by what some applications wanted.”*

When asking participants why they chose the app they actually downloaded instead of another, price is mentioned more often as in the former description of their decision-making process. Hence, the factor price and whether the app can be downloaded for free seems to play a major role in the decision-making process. Some participants simply considered an app as feasible to download *“because it is free”* or they generally *“never”* pay for an app. Some participants indicated a willingness to only pay under certain circumstances or for specific types of apps as they prefer *“not to give my banking information for a simple game.”* Free alternatives available in app stores was also a reason for not paying for an app, or as one participant put it: *“why would I pay for something that I can get for free?”* Other participants indicated that they are only willing to pay for an app if someone within their social circle recommended it: *“only when good friends advise on a paid app, I will buy that straight away.”* Other users would like to test an app before buying it because *“if I have to pay for every app I test, and most probably discard, that’s not a good incentive for the app writer to make it better.”* Again, design of an app, functionality and requested permissions seem to have top priority in the evaluation process as well as 9%, 7% and 6% of the sample mention these factors respectively. The quotes of these categories are similar to the ones mentioned in the description of the decision-making process and will therefore not be presented twice.

3.4.5 Privacy and security related considerations in the reviews

When looking at the reviews participants wrote about their app, it is remarkable that privacy and security considerations were barely mentioned. In line with the descriptions of the decision-making process, factors that are visible and that can be directly experienced by the user were dominant in the reviews. An analysis of the reviews showed that usability (19%), ease of use (17%), design (16%), and functionality of the app (12%) were most often mentioned. Privacy and security issues only played a minor role, if any, in the participants’ evaluations of the chosen apps for others. Although the requested permissions had played a minor role during the downloading and installation process,

they were virtually absent in the reviews. Privacy and security considerations were discussed in only 8% of the reviews. Participants, for instance, were worried about the permissions requested, because the app *“asked for too many permissions like creating and deleting accounts and even changing the password of accounts.”* One participant linked permissions to the functionality of the app as he recognized that *“the app asks for permissions that are (in my opinion) not needed for normal use of the app.”* Consequently, this participant declined the permissions after having downloaded the app. Another participant reported having changed the permissions afterwards because he felt *“that the game required too many permissions (i.e. wanting to access your contact list)”* and he *“turned the permissions off immediately.”* However, participants mentioning privacy concerns still had chosen to download the app that was deemed to have privacy problems.

3.5 Discussion

The purpose of this paper was to research the phenomenon of the privacy paradox. More specifically, this study aimed at exploring whether or not the privacy paradox is in fact observable in actual behavior and not attributed to a given intention/attitude-behavior gap as reported in literature (Baek, 2014; Barth & De Jong, 2017; Dienlin & Trepte, 2014).

First, by means of an experiment, actual behavior was examined with regard to downloading and installation patterns. Here, we controlled for technical expertise and financial considerations by studying a tech-savvy user group and providing monetary compensation. Second, we determined the factors that play a role in the decision-making process as it pertains to selecting, downloading and installing an app at two moments in time: before and after the actual installation process with a major focus on privacy and security. The reasoning behind this model was to compare declared intention and attitude with self-reports on actual behavior. By doing so, data from the experimental part of this study could be collated with the results of self-reports. Both parts of the study served for confirmation and validation of each other. In this final section, the main results will be reviewed and implications for user empowerment, support and future research will be given.

3.5.1 Main findings

The results showed that, in general, users do not rank privacy and security related aspects as a high priority when considering factors that guide the downloading

and installation process of an app. Factors such as price, ratings and design seem to play a major role in the downloading decision, although participants indicated previously that usefulness, functionality and trust in particular influence their app choice. This is in line with the findings of Kelley et al. (2013) who found in their study on app decision-making process that users consider factors such as cost, functionality, design, ratings, reviews and downloads as more important than requested permissions. Apps with higher ratings are higher listed in the display search function of the app-store. Hence, these apps will be more often recognized, valued and downloaded by users (Dehling et al., 2015). Furthermore, while self-reporting on the actual downloading and installation process of an app, participants mentioned that permissions play a major role in their selection of an app. However, the experiment showed that participants did not behave in accordance with their previous indications. When talking about location based applications, Zafeiropoulou et al. (2013) found that users do not act according to their previous stated privacy preferences as benefits of such apps outweigh privacy concerns. This behavior is also represented in the peer reviews, written by the participants, as privacy and security related factors are rarely mentioned. This leads us to assume that privacy and security as concepts seem to play a minor role in the mental representations of users. Functionality and design seem to outweigh privacy concerns or privacy is not considered as an important part in the overall evaluation or recommendation of an app. Hence, it seems that privacy does not play a role in the social representation of factors that are considered important when talking about, reviewing, or actually downloading an app. This is in line with Kehr et al. (2015) who suggest that the social representation about privacy is not yet formed by lay-users. It would seem that privacy is but a marginal note in the adoption process of mobile apps. This prompts the question: How much do consumers really value their data and privacy?

3.5.2 Implications

The results of this study were ascertained from a tech-savvy target group possessing technical expertise above the average user. Not only did participants describe themselves as informed about technical issues regarding smartphones and mobile applications, but the sample expressed privacy concerns by implementing data protection measures and restricting unauthorized information distribution to third parties. Despite these concerns, these individuals were not willing to pay for an app that asked for less information by means of permissions. These findings correspond with the study of Williams et al. (2017) who found that knowledge about risks regarding IoT does not prevent users from buying and using such products. Although users knew about certain risks pertaining

to IoT products and revealing private information, such users seemed to value their personal data less and struggled significantly more to protect their data compared to non-IoT users. Furthermore, the paradoxical behavior (having concerns and revealing private information) was more present among IoT users compared to non-IoT users. This again raises the question about privacy valuation of mobile app users and how they actually perceive privacy aspects in an online environment.

On a similar note, this paradoxical behavior prompts the question as to whether or not even subjects with a technical background understand enough about permissions and their potential ramifications. The understanding of permissions is a crucial component in privacy consideration as without sufficient knowledge about permissions warning, users are unable to make adequate privacy decisions (Felt et al., 2012). Benton et al. (2013) also concluded that permissions are difficult to understand for users and ineffective when it comes to privacy considerations. In their study on user attention, comprehension and behavior regarding Android permissions, Felt et al. (2012) determined that only 17% of participants paid attention to permissions and almost half of the sample did not notice permissions at all. Comprehension levels were also rather low. Deuker (2010) concluded that enhancing privacy awareness alone is not the solution for this complex problem. Users need supporting tools in order to react according to their privacy preferences.

When looking at even the newest version of the Android permission system, permission requests take place rather late in the process. Users have already installed the app on their smartphones before being asked to grant certain permissions. One might hypothesize that this is a conscious ploy that takes advantage of a weaknesses in human nature: If a user has already carried out the cognitively demanding efforts employed when selecting an app such as evaluating price, design and reviews, then downloads and installs an app, the cognitive effort to evaluate the ramifications of granting permissions afterwards might be too high, leading to a situation in which security risks are accepted despite privacy concerns. Furthermore, explanations of single permissions (or from Android 6.0 on the higher permissions groups) are not directly visible for the user, meaning they have to perform a further task to obtain this information. This might lead to a decision guided by simple heuristics: the user denies and uninstalls an app, implying that all efforts made are nullified, or the problem is relativized and accepted despite privacy concerns. The latter seems to be more common and could prove a major factor in the privacy paradox. This assumption however, has not been researched to date. The scope of permissions and their implications might be considered an inherent part of privacy valuation. As such,

more research is needed into this domain, not only from social and behavioral perspectives, but from a technical point of view as well. The key questions remain the same for both perspectives: How can we improve the existing Android permission system to enhance user empowerment and what are the alternatives to the permission system in question? Privacy has to become part of the social presentation of users in order to be a valuable asset. Understanding the public's mindset as it applies to the handling of mobile applications and adapting interface design (of permissions) to cognitive styles, raising awareness, increasing knowledge and providing support while simultaneously empowering users seems to be the key to putting online privacy on the agenda. Considering the impact of privacy and security risks, Dehling et al. (2015) suggest that the type of app (e.g., what and how much data are gathered) should be taken into account for protection mechanisms but also for tailored risk sensitization (e.g., what is the impact of a specific data sharing process). However, empowerment of users to make self-determined and self-protective decisions with regard to their personal data behavior should be a top priority. Promotion of privacy protection behavior should help users to overcome the paradoxical behavior that we can still observe when dealing with mobile computing and in particular with apps.

3.5.3 Limitations and future research

This study reflects a first attempt to put the existence of the privacy paradox to the test in more extreme circumstances, focusing on actual behavior instead of behavioral intentions and eliminating some of the potential explanations for the paradoxical behavior (a lack of knowledge, privacy awareness, or money). The results support the existence of the privacy paradox in these circumstances. However, it is important to keep in mind three limitations of our study. The first limitation involves the sample size. Mainly due to the intensive nature of the data collection, our sample size was rather small, even after two years of collecting data. The experimental results of our study are quite clear, even within the limited sample, but a larger sample would have enabled us to further explore the relationships between knowledge, awareness and download considerations. Future research could study such relationships more extensively.

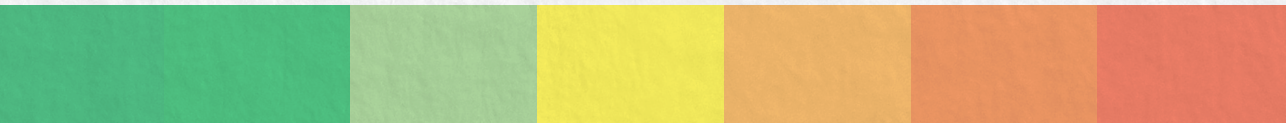
A second limitation involves the artificial context used in the research. The participants in our study were only required to download and use the app for one week, and we cannot be certain whether they would have downloaded and used the particular type of app in real life. It is imaginable that the participants acted somewhat pragmatically, knowing that they could discard the app one week later. However, during that week, they still exposed themselves to the privacy and security risks of the app they downloaded. Future research could focus more

on apps that users actively use for a longer period of time. For instance, by first inventorying and discussing the apps users have on their mobile phone, then analyzing the privacy and security issues of these apps, and discussing them in a second round with the users.

A third limitation involves the types of apps used in this study. We selected two types of apps, as typical representatives of apps with a utilitarian or hedonic purpose. However, the privacy and security considerations of users might be totally different with other types of apps. It is imaginable that users' privacy and security considerations are more important when, for instance, the app focuses explicitly on their health status, their social network, their geographical location, or their photos, videos and sound recordings. Future research could therefore focus on users' download behavior and considerations regarding different types of apps.

3.5.4 Conclusions

The main purpose of this study was to examine whether or not a given paradoxical behavior is still observable in users not disadvantaged by a lack of technical knowledge, privacy awareness, or financial means. We can indeed confirm, that despite the fact users still claim to be concerned about the potential misuse of their personal data, they remain unwilling to invest either the time and effort or the money necessary to protect their privacy. Despite their technical backgrounds and a higher than average understanding of privacy intrusion possibilities, participants were not willing to pay for their privacy. We highly question whether or not privacy as a concept is already implanted in users' perception and social representation. If not, this might explain the discrepancy between claims to highly value privacy and the behavior that indicates otherwise.



4

Lost in privacy? Online privacy from a cybersecurity expert perspective

Barth, S., De Jong, M. D. T., & Junger, M. (2020). *Lost in privacy? Online privacy from a cybersecurity expert perspective*. Manuscript submitted for publication.

4.1 Introduction

Smartphones are omnipresent and have become an integral part of our daily lives. Applications ('apps') running on mobile devices support and enable many of our daily activities, including information provision, entertainment, amusement, social interactivity, biological monitoring and online shopping. By installing and using apps, users generate vast amounts of personal data that may or may not be passed along to other parties. The vast majority of this data collection is not transparent and therefore difficult to understand. A recent study of the Pew Research Center showed that 79% of the online service users have concerns about their personal data (Auxier et al., 2019). Despite these concerns, the number of app downloads continues to increase and is expected to exceed 250 billion by 2022 (Statista, 2019). This suggests that any privacy concerns users might have are not significantly influencing their decisions to install and use apps. In other words: potential privacy infringement issues are taking a back seat to satisfy their immediate practical, social, informational or entertainment needs/desires. The immediate benefits of downloading and using a particular app are prominent and the possibility of privacy-compromising consequences in the future is easily accepted (Kehr et al., 2014; Pentina et al., 2016; Shklovski et al., 2014).

The discrepancy between privacy concerns and actual behavior is known as the privacy paradox: People say that they care about their online privacy but nevertheless disclose personal and sensitive information without hesitation (Acquisti, 2004; Barnes, 2006; Barth & De Jong, 2017). Most of our insights on privacy-related attitudes and behaviors are based on research with general users¹, suggesting that a lack of knowledge might be an important factor. It poses the question as to whether or not understanding the exact privacy risks involved in downloading and using apps leads to a heightened privacy awareness that in turn affects behavior accordingly. Interestingly, a recent study by Barth et al. (2019) suggested that more knowledgeable and privacy-aware users—specifically, advanced computer science students—subject themselves to practically the same privacy risks as general users. When it comes to instant messaging for example, De Luca et al. (2016) also found that experts engage in the same unsafe behaviors as 'less digitally literate' users. Although our insights regarding cybersecurity and privacy are predominantly based on the research and development work of cybersecurity and privacy experts, there has been little research to date on their personal views and behaviors regarding online privacy.

1 Within the realm of this dissertation, the definition of a general (or 'lay') user is a user who does not possess any specialized online privacy and/or cybersecurity expertise.

In this chapter, we present a qualitative interview study investigating how privacy and cybersecurity experts (hereafter ‘experts’) deal with online privacy on their own smartphones. We addressed the following two research questions:

RQ1: How do privacy and security experts value their personal online privacy?

RQ2: How do privacy and security experts evaluate and use mobile apps?

4.2 Earlier research

With the rise of Information and Communications Technology (ICT), online privacy became an integral part of the interaction between humans and technology. It became more important than ever that users define, preserve and monitor their personal boundaries when it comes to the disclosure of data. With smartphones and apps, the privacy discussion became even more prominent. As smartphones combine sensors, 24/7 connectivity, real-time tracking, data aggregation, and profiling with simple and limited interfaces, their lack of transparency and potential for personal intrusion is far greater than that of desktop computers and notebooks. Inexperienced users can easily lose their way in the privacy jungle and require support when making informed decisions on their valuation of privacy.

4.2.1 Experience of privacy: A user perspective

Many studies revealed that users do not translate their general concerns about online privacy into actual behavior, a contradiction known as the privacy paradox. The privacy paradox seems to be particularly applicable to smartphone contexts (Benenson et al., 2012). Various explanations for the privacy paradox have been provided: (1) privacy threats and benefits are rationally weighed, whereby benefits outweigh privacy threats; (2) privacy threats and benefits are weighed, but the outcome is skewed by irrational factors or bounded rationality; and (3) privacy threats are not even included in users’ considerations (Barth & De Jong, 2017).

In the context of e-mail encryption, Renaud et al. (2014) used the metaphor of a staircase to illustrate the stages users have to go through regarding the protection of their online privacy. The first step is privacy awareness: knowing that privacy may be an important issue in online activities. The second step involves privacy concerns: worrying about their own online privacy. This is followed by a full understanding of privacy threats – and one step higher, recognition of the need to actively protect their online privacy. The next steps are knowing how to protect themselves and being able to do so respectively. The top

step—not being side-tracked—accounts for irrational elements that may interfere. While interviews with users confirmed the first five steps, users did not reach the step in which usability of protection measures became an issue.

Although there is no consensus about the underlying mechanisms of the privacy paradox, it seems plausible that the human mind is bounded by nature and cognitive overload makes rational decision-making highly unlikely (Simon, 1982; Veltri & Ivchenko, 2017). One can think of the mechanism of delay discounting, which refers to the tendency that outcomes that are remote in time have less impact than immediate outcomes (Odum, 2011). Long-term privacy threats are hard to estimate (Shklovski et al., 2014), while the costs of paying for secure and safe apps are obvious and immediate. This is compounded by the fact that the abuse of personal data is often difficult to determine and data breaches are considered relatively unlikely.

Users might adjust their attitudes toward the value of their private data for reasons of pragmatism, as the wish to use apps seems to be stronger than the conceived risks of data misuse (Debatin et al., 2009). They may for instance, see themselves as not significant enough a target for potential fraud attacks and therefore underestimate the risk probability, shift responsibilities to other parties such as the app store or the government (Volkamer et al., 2015), or erroneously assume to be less vulnerable to privacy threats than other people (Debatin et al., 2009). A users' desire for any particular app is directly related to their willingness to disclose personal information when acquiring and using such. This may particularly apply to social networking apps, due to group dynamics (Taddicken, 2014) and their significance and ubiquity in people's lives (Debatin et al., 2009).

An additional important consideration is that processes running in the background of mobile apps are difficult to understand for users (Acquisti et al., 2016). Unexpected data flows resulting from conscious transactions of personal data with other parties are particularly ambiguous for users, although strongly affecting users' privacy (e.g., the collection of meta-data for profiling; Bräunlich et al., 2020).

Two other research findings further complicate users' privacy-related behaviors with mobile apps. Firstly, research showed that privacy and security risks are perceived to be more likely on desktop computers than on mobile phones (Volkamer et al., 2015). Secondly, Choi et al. (2018) showed that privacy fatigue increases disengagement in protective measures and ostensible indifference toward privacy violation, eventually outweighing privacy concerns.

4.2.2 Understanding of privacy: The role of expertise

Concisely defining privacy in all of its applicable and/or relevant levels is by no means a simple task, especially in online contexts. Its intangible and non-urgent nature can at least partially be attributed to a lack of knowledge among users about the risks of online behaviors and the consequences of data disclosure (Bandara et al., 2017). This might imply that technical expertise leads to different evaluations of personal data and precautions online that deviate from the norm. In a study on phone embedded tracking, Ketelaar and Van Balen (2018) found some paradoxical evidence for this assumption: Smartphone users with considerable knowledge about the technical mechanisms behind data collection and types of data targeted had lower privacy concerns and more positive attitudes toward location tracking, likely attributable to their ability to take better precautions against data gathering.

In a similar vein, knowledge of and attitudes toward security measures might explain that personal data are protected differently and eventually more effectively by experts than by non-experts (Ion et al., 2015). However, in their study on attitudes toward instant messaging services, De Luca et al. (2016) found that despite their technical knowledge, experts showed the similar risky behaviors online as general users. Interestingly, Reidenberg et al. (2015) showed that the privacy policies of web services are especially ambiguous when making statements about data sharing and sensitive information. This ambiguity prompts misinterpretations about data sharing practices that might lead to the unintentional disclosure of personal data, an effect that is not only observable among general users, but also among privacy policy experts as well. Still, general users seem to be more prone to decisions based on false security and privacy assumptions than experts. Similarly, the mental models general users have of Tor architecture are found to be superficial, incomplete, and rather abstract compared to those of experts, who showed a deeper understanding of threat models. The false assumptions about security made by general users may have serious impacts on their privacy-related behaviors online. Interestingly, experts have knowledge gaps as well, but less serious than novice users (Gallagher et al., 2017). Still, their knowledge gaps appear to have similar effects on their privacy-related behaviors. From these earlier findings, one might infer that even considerable technical knowledge can be overridden by situational or external factors (e.g., time constraints, money considerations, group pressure) and that being technically literate does not automatically lead to different evaluations of privacy threats or deviations from the 'norm' as they pertain to precautions online.

4.2.3 Estimation of privacy: Proxies that guide decision-making

The aforementioned research findings underline that, within the complex and highly technical mobile environment and the boundaries of human cognition, it is hard for general users to make informed decisions about the online privacy and security of apps and that this may also apply to users with considerable technical knowledge. Users must rely on several signals to decide whether to download an app or not. Permission requests could function as a proxy for privacy-related information. Still, less tech-literate users often encounter difficulties understanding them and click through such prompts without paying much attention. As a result, factors such as app design, ratings, download rates, reviews, costs, functionality and peer recommendations may easily outweigh privacy-related considerations, or privacy is not even considered in the evaluation and adoption process of apps (Benton et al., 2013; Chin et al., 2012; Felt et al., 2012; Kehr et al., 2015; Kelley et al., 2012). The app selection process seems to be guided by a ‘take the first’ heuristics (predominantly based on star ratings), a recognition heuristic (e.g., hearsay or prior experiences), or a vote heuristic (based on ratings and reviews). Furthermore, apps for free seems to be a guiding principle (Dogruel et al., 2015; Joeckel et al., 2017). Technical knowledge and considerations of permissions as proxies for privacy risks seemed to be outweighed by ratings, app design and costs (Barth et al., 2019).

A more sophisticated understanding of the internet and its underlying (technical) mechanisms does not automatically result in a significant deviation in privacy protection actions. Experts’ willingness to take protective precautions may still be overruled by internal considerations (e.g., a nothing-to-hide attitude) or contextual cues such as familiarity with the service or company, or symbols indicating privacy protection (Kang et al., 2015). Experts and general users might evaluate privacy threats differently and use different proxies to guide decisions on adopting an app. For instance, experts claim to take permission information and reviews predominantly into account, whereas general users claim to rely more on app descriptions and ratings (Jorgensen et al., 2015). However, the question whether or not technical knowledge leads to a more sophisticated evaluation of apps and different privacy-related behaviors remains unanswered.

4.3 Method

To answer the research questions, we conducted a qualitative study based on semi-structured interviews with privacy and cybersecurity experts. The interviews served multiple purposes. In this chapter, we focused on the experts' personal views as they pertain to online privacy and online behavior. In another article, we will report on their professional knowledge about online privacy. The study was approved by the ethical committee of the BMS faculty, University of Twente. Below, we will outline the details of the research.

4.3.1 Participants

A total of 40 privacy and cybersecurity experts from our academic and professional network were invited by e-mail to participate in this study. Inclusion criteria were a job description focusing on privacy and cybersecurity, coupled with mobile app and online privacy experience. In the invitation e-mail, the potential participants were provided with a summary of the research objectives and a preview of the interview questions, so that they could decide for themselves whether or not they would be suitable as participants. Eventually, 20 participants were personally interviewed (one of whom withdrew afterwards) and one participant remotely, resulting in a final sample of 20 privacy and cybersecurity experts. Participants received no compensation for taking part in this study.

The participants were aged between 24 and 54 (mean = 38.1 years). All participants were male with engineering backgrounds. Their most frequent job descriptions were researcher, scientist, or analyst ($n = 9$), followed by security, technical, or program manager/coordinator ($n = 5$), consultant ($n = 3$), developer ($n = 2$), and software engineer ($n = 1$). All participants worked for public or private organizations operating in the domain of cybersecurity in The Netherlands.

4.3.2 Research instrument

The semi-structured interviews were based on an interview guide with open-ended questions. First, participants were asked about their background, specifically their age, education, and current job. Participants were then asked to provide an impression of their job activities and responsibilities, as well as elaborate on the role of privacy and security issues in their jobs. For the first research question—how experts value their personal online privacy—we asked the participants about the importance they assign online privacy in general and to mobile apps in particular. For the second research question—how experts evaluate and use mobile apps—we asked them about their personal smartphone

usage and their practices regarding downloading apps (e.g., the cues they use to guide decisions about downloading an app).

4.3.3 Procedure

Each interview lasted between 45 and 90 minutes. Before starting with the semi-structured interviews, information about the research aim was shared. All participants signed an informed consent form and agreed to the interview audio being recorded. After that, questions from the interview guide were discussed with participants, giving them sufficient possibility to present their ideas on the topic in question. At the end of the session, the participants were debriefed and thanked for their participation.

4.3.4 Analysis

Interviews were transcribed verbatim and any personal information that could identify participants was removed. After that, the data were imported into ATLAS.ti for code creation and analysis. Starting with an open coding procedure, a list of codes was derived based on research questions, literature and reading through all of the transcripts. Meaningful text passages were highlighted and codes were attached to them until a point of saturation was reached. The units of analysis varied from single buzzwords to statements made in multiple sentences. The list of codes was discussed with the co-authors, considerably shortened, which eventually resulted in the following main code categories: (a) age and working position, (b) views on privacy, (c) review of signals, and (d) mobile phone usage. To assess the reliability of the codebook, 10% of the sample was coded by two independent coders (the first author and an independent researcher). The procedure was repeated twice, with the codebook being refined after each round and intensive discussions between both coders. Eventually, Cohen's kappa was .75, indicating substantial agreement. The remaining 90% was then coded with the revised codebook by the first author.

4.4 Results

4.4.1 Value of online privacy

The results show that the participants differed in their opinions about the value of privacy, eventually resulting in three groups: (1) experts who are concerned about their privacy ($n = 7$), (2) experts who are conscious about their privacy but not overly concerned ($n = 7$), and (3) experts who do not pay much attention to their privacy ($n = 6$). These groups correspond to the three user

categories distinguished by Westin's (1967) privacy orientation index: (1) *privacy fundamentalists*, who are "at the maximum extreme of privacy concern...[and] are the most protective of their privacy," (2) *privacy pragmatists*, who "weigh the potential pros and cons of sharing information...after this, they decide whether it makes sense for them to share their personal information," and (3) *privacy unconcerned*, who are "the least protective of their privacy – they feel that the benefits they may receive...far outweigh the potential abuses of this information" (cf. Kumaraguru & Carnor, 2005, p. 15). Table 4.1 summarizes the three categories, along with typical quotes regarding online privacy value and mobile phone behavior. We will use these categories to provide in-depth descriptions of participants' views and behaviors.

Table 4.1 Expert groups based on Westin's Privacy Orientation Index (POI), with illustrative quotes

Privacy orientation	n	Example quotations	
		Value of online privacy	Mobile phone behavior
Privacy fundamentalists	7	"...on an Android phone, Google tracks your data by default. This can be what shops I visit, what hotels I book. The only thing I did is to switch on my mobile phone and immediately data gathering starts. I perceive this as an invasion of my privacy, because Google should not do that kind of things....Yes, privacy is really important to me."	"I think that this is dangerous. I avoid such things...if an app does not need to send SMS, why should I give permission to that. I don't trust it then. Sorry."
Privacy pragmatists	7	"I always try to figure out how they apply security measures and how they handle data. These kinds of things. And what is the risk that you might take. And then I try to weigh risks against benefits. But do I really worry about it, no, I cannot say that. Yes, of course, there are certain risks when you do things online. I know that...but..."	"I know quite well who knows what about me. And I don't worry about it because I only tell people, Facebook or Google for instance, about things I want them to know. Yes, of course, some organizations have access to my personal pictures, for instance. Therefore, I have only pictures on my mobile phone they are allowed to see...but privacy is not very important to me. And this reflected in the fact that I install everything on my mobile phone."
Privacy unconcerned	6	"You know the risks...but I have nothing to hide, so I don't worry about it."	"There are no apps I worry about."

4.4.1.1 Privacy fundamentalists

A total of seven of the 20 participants found their online privacy important or very important. Five of them claimed that their personal views on privacy matched their professional views. As a matter of fact, their strong views on the importance of privacy were a reason for several of them to find privacy and security-related jobs. One participant stated that privacy for him was a moral decision. Using an online service is a “*transaction that takes place at a given moment in time, but beyond that transaction, the service is not allowed to obtain information.*” Other participants, however, mentioned a reverse relationship, with their professional knowledge about privacy and security influencing the amount of attention they pay to privacy issues in private settings. One participant argued that deciding about privacy in his job is easier than doing so in real-life situations: At work he can follow clear definitions of privacy, but in his private life he is confronted with continuous decision-making within changing contexts.

Although these participants seemed to be very conscious about their personal privacy and aware of privacy risks, they tended to use online services nevertheless, including the more risky ones. They justified this behavior referring to external factors such as time constraints or group pressure, but also to internal factors such as complacency, convenience, or an irresistible desire to use a certain app: “*I don’t want apps to disclose my personal data to others, although I know that this happens...and yes, I use them nevertheless. For the mere reason that others with whom I communicate, use that app.*” Although privacy was high on their agenda, they sometimes made decisions of which they knew “*that they are not optimal,*” because the urgent need to own and use an app outweighed their concerns. Such justifications indicate that a given discrepancy between the stated privacy concerns and actual behavior was sometimes present. In addition, it mattered to these participants if a service could be held legally accountable for data misuse and privacy violations.

To minimize risks and resolve cognitive dissonance when using apps, the participants said they evaluated whether or not the app permissions required remained within the realms of their personal privacy boundaries and corresponded to the core functionality of the apps. If this was not the case, permissions were seen as “*dangerous*” and were sometimes denied. Some permissions though, were considered hard to interpret and fathoming them and all of their implications was considered demanding by some participants. However, one participant made a contradictory statement about the importance of understanding permissions, relativizing the impact of understanding troublesome permissions: “*In most cases I look at the permissions, try to*

understand and explain them....However, I have never denied an app because of the permissions so far.”

Despite the fact that some participants argued that the selection and use of apps occurred very consciously and that all measures to protect personal data were taken, they were unable to clearly describe their risk assessment process. App usage was kept to the minimum necessary, but “*standard apps like YouTube, Twitter, Facebook*’ are used, *‘even though preferably not, but this is a necessary evil.’*” However, if an app is perceived as absolutely untrustworthy, it will not be downloaded. In most cases, the participants expressed a certain degree of mistrust as it pertains to permissions.

4.4.1.2 Privacy pragmatists

Seven other participants could be characterized as privacy-conscious but not overly concerned. They generally applied rough risk-to-benefit calculations before downloading and using mobile apps. While they were quite aware about certain threats to their privacy, they considered their personal privacy to be of less significance than their work-related privacy. Professionally, the right to privacy was very important to them, but this did not manifest itself as clearly on the personal level. They emphasized that users should always be able to choose which data to disclose when using online services, but appeared to pay less attention to the data they disclosed themselves. They were fully aware that some companies know a lot about mobile phone users. But as long as they were not directly confronted with it—or disadvantaged by it—the data gathering practices of companies were considered to be acceptable.

Similar to the group of privacy fundamentalists, these participants found it difficult to assess the risks of apps. Some participants tried to remedy such difficulties by limiting the online services or apps they actually used without restricting data gathering. One participant stated:

“Yes, on the one hand, I consciously decided to use Google products, but Google products exclusively, so that one big company knows much about me, I don’t think that’s something really bad. I don’t worry about the information they have about me, but I do not want them to tell me point blank that they have that information about me.”

Furthermore, some participants questioned the control someone actually can have over personal data in online environments. With high effort and considerable knowledge about technical issues, one might gain control of personal data. Still, in many situations, users have neither the time nor energy

to engage in protective measures. The interviews suggest that these participants are well aware that personal data are not always treated confidentially but are willing to use such online services nonetheless. Because of their backgrounds in privacy and cybersecurity, they admitted that they should pay more attention to online privacy and the disclosure of personal data than they actually did. One participant pointed out that his offline life corresponds with his online life and that more or less the same information about him is available in both contexts. At the same time however, this participant was aware that the business models of online services are not comparable with those of offline services, arguing that this is “*not a big deal.*”

Several contradictory statements were made in this group of participants in particular. They appeared well aware that their data were not always treated properly and that their personal privacy might be infringed. Still they tended to use online services without significantly protecting themselves from privacy threats. Interestingly, some participants called themselves naive, as they could not imagine that “*big brother is watching you*” all of the time. Plausible strategies to act in line with their privacy concerns would include being very selective as it pertains to which data to disclose and which data to keep private, or to rely upon paid apps rather than free ones with questionable privacy privileges. However, it appeared that neither of the strategies were substantially implemented by the participants.

4.4.1.3 *Unconcerned users*

The last group of six participants can be described as indifferent to their personal privacy. Most of these participants claimed to know the risks but have no qualms about disclosing personal data online. In this group, the “*I have nothing to hide*” and the “*online privacy does not exist anymore*” arguments were mentioned often. Furthermore, the participants had no objections to their data being analyzed generically, as long as the individual remains anonymous. To them, privacy seemed to be manageable at first sight, but became complicated and nuanced after obtaining deeper insights into data gathering practices. However, this realization did not lead to a change in online behavior but to “*I should have known better*” resignations.

Unconcerned participants were aware that they have to give something in return for using certain online services, but still perceived the benefits to weigh more than their perceived level of lost personal privacy. Furthermore, they questioned the availability of viable alternatives. One participant stated:

“Of course that might trigger an inner alarm, but what is the alternative to the permissions they ask for? Yes, you could decide to not install that app eventually. But the question remains, do you choose to play safe or to not install the app? I don’t think that I won’t install the app because of the permissions they ask for....Of course, it’s a dilemma because it costs me something.”

Mistrust of certain services remained, but risks were accepted because of a lack of alternatives and a strong wish to own and use an app. This strategy of neglecting privacy risks was followed so long as nothing serious happened with their data. Furthermore, peers were taken as a reference point (*“If they do, it should be ok”*). The organization behind the app was also used as an important consideration (*“All my apps are from more or less big companies. That’s helping me to trust them. Maybe, that’s a bit naive but...”*). Moreover, privacy cynicism seemed to manifest itself in participants’ reasoning regarding online privacy, starting with feelings of helplessness (*“I am unable to change the situation anyway”*) and eventually leading to an acceptance of data handling practices, even if they are not in line with one’s personal views.

Although indifferent users seemed to pay little attention to their privacy, it seemed that privacy boundaries, as far as possible, still played a role in their considerations (*“I have nothing to hide...But anyway, I try to consider the purpose....And yes, I try to choose a serious app that aligns with it”*).

4.4.2 Review of privacy-related cues

Despite the differences described above, 90% of all participants indicated to review privacy-related cues before downloading an app. The cues functioned as proxies to evaluate whether or not an app is deemed trustworthy enough to download. Most participants generally made a superficial scan of the app instead of conducting any real risk analysis. When looking at the group segmentation of fundamentalists, pragmatists and unconcerned users, one might expect that fundamentalists would evaluate the most privacy-related cues when considering an app. This was not the case: The average number of signals used was comparable in the three groups (2.4, 2.3, and 2.5 cues, respectively).

Table 4.2 gives an overview of the cues considered by the participants. Of all the cues experts mentioned, requested permissions were given priority. However, it is unclear whether the requested permissions really influenced participants’ decisions as to whether or not to download an app. Only a few of the participants stated that they had in fact decided against downloading an app based on its disproportionate permission requests. The majority reviewed permissions, but

as long as they—at least to some degree—were related to the app’s functionality, they downloaded the app despite the fact that not all of permissions were deemed entirely legitimate. The second most considered cue were reviews, in which other users share their experiences with the app. Participants admitted that they often only considered the most positive and negative reviews, or those most easily accessed by positioning. It should be noted however, that the participants admitted during the interviews that they paid no attention to the credibility of the reviews.

Interestingly, 35% of the participants considered the developer or owner of the app as an important cue for trustworthiness. They wanted to know where the app came from, whether or not the developer of the app is officially registered, and/or whether or not the developer or company behind the app had already published other apps. One participant mentioned that the ‘certified developer flag,’ a recommendation of Google, helped him to verify the trustworthiness of the app. Another verification process was to assess the number of app downloads along the lines of: the more downloads, the better the app is and therefore, the more trustworthy it is. Furthermore, 25% of the participants looked at the description of the app or the ratings displayed in the App Store. In the descriptions, professional language seemed to be important, as unprofessional or plain language use was equated with less trustworthiness. Apps with bad or no ratings were disqualified from further consideration. Other cues mentioned by individual participants were screenshots of the app in the app store, word-of-mouth, the number of app versions/updates and familiarity.

In conclusion, only two participants did not look at cues at all, simply downloading everything they were interested in, willing to chance what the future would hold. Sometimes, they uninstalled an app afterwards, citing usage and/or trust issues that had arisen. The other participants took at least one cue into account before they actually made decisions on which apps to download on their phones. However, the extent of their considerations differed between participants.

Table 4.2 Cues reviewed by experts before downloading a mobile app

Expert	Privacy Orientation	Permissions	Reviews	App developer/owner	Downloads	App description	Ratings	Total
1	Fundamentalist	x	x	x		x		4
2	Fundamentalist		x	x	x			3
3	Fundamentalist	x						1
4	Fundamentalist	x		x				2
5	Fundamentalist	x		x	x			3
6	Fundamentalist		x	x				3
7	Fundamentalist	x						1
8	Pragmatist		x		x	x		3
9	Pragmatist	x						1
10	Pragmatist							-
11	Pragmatist	x	x		x	x	x	5
12	Pragmatist	x						1
13	Pragmatist	x	x		x			4
14	Pragmatist	x		x				2
15	Unconcerned	x					x	1
16	Unconcerned	x			x			2
17	Unconcerned		x	x		x	x	4
18	Unconcerned							-
19	Unconcerned	x	x				x	3
20	Unconcerned	x	x		x	x	x	5
Total		14 (70%)	9 (45%)	7 (35%)	7 (35%)	5 (25%)	5 (25%)	

4.5 Discussion

4.5.1 Main findings and theoretical contribution

Our interviews with experts from the privacy and cybersecurity field can be summarized into three main findings that contribute to our understanding of online privacy: (1) technical knowledge does not automatically lead to more privacy-conscious behaviors, (2) experts and general users do not differ in the way they justify risky online behaviors and (3) although experts and general users utilize different cues to evaluate the appropriateness of apps, the resulting online behaviors are largely similar.

4.5.1.1 The role of expertise in valuing privacy and protecting personal data

The segmentation of experts into groups corresponding with Westin's (1967) privacy orientation index showed that technical knowledge pertaining to privacy and cybersecurity does not automatically lead to a higher privacy evaluation or more precautions to protect personal data online. On the contrary, experts may engage in different data handling practices that have similar outcomes: (1) a fundamentalistic view on privacy *but* still vulnerable to biases, (2) a pragmatic view on privacy *and* vulnerable to biases and (3) an unconcerned view on privacy and little or no consideration of risks. Whatever importance experts attached to their online privacy, the strategies they used often led to unsafe, sometimes careless, online behaviors. Irrespective of their privacy orientation, the experts, just like general users, struggled with the strong, immediate temptations of free apps and had mechanisms in place to temporarily relativize the importance of privacy to justify questionable judgment.

As such, the privacy paradox applies to the experts' attitudes and self-reported behaviors regarding mobile apps as much as it applies to those of general users. Based primarily on studies with general users, research on the privacy paradox (Acquisti, 2004; Barth & De Jong, 2017) shows that, irrespective of the value users attach to their online privacy, data disclosure happens rather easily, especially in the fast changing environment of mobile computing (Kehr et al., 2014). General users are often unable to resist downloading apps and giving up personal data because of biases in human decision-making—for instance, too many complex factors to be considered, time constraints, or optimism bias (Acquisti, 2004; Shklovski et al., 2014). However, it is also argued that technical expertise might, at least to some extent, compensate for such biases. Technologically savvy users are assumed to be able to protect their personal data more effectively than general users (Ion et al., 2015; Ketelaar & Van Balen, 2018) or at the very least, technical understanding brings more

clarity to complex online environments (Bandara et al., 2017). The assumption that experts who are specialized in privacy and cybersecurity are better able to evaluate and act upon the potential risks of data disclosure was not confirmed by our interviews: Similar to earlier findings of Barth et al. (2019) and De Luca et al. (2016), the self-reported online behavior of users with higher technical expertise closely resembled that of less knowledgeable users.

4.5.1.2 Justifications of risky online behavior

In addition to showing that the online behavior of experts appears as unconcerned as that of general users, our interviews also revealed that the arguments they provide to justify their risky online behaviors are comparable to those of general users as well. First, experts experience a lack of alternatives. For instance, if there is only one suitable app available, they may be inclined to accept everything, even if it might affect their privacy negatively. A sense of being overwhelmed and learned helplessness in an online environment leads to data disclosure, despite privacy concerns and an inherent wish to protect private data (Hoffmann et al., 2016). Several experts claimed that the individual user resigns to being unable to change data handling practices anyway.

Second, and similar to the perceptions of general users, statements associated with time constraints (Flender & Müller, 2012) and a disregard of potential long-term effects (Acquisti, 2004), diminished the necessity for personal privacy protection. Experts indicated having neither the time nor willingness to carefully consider every single aspect that might infringe upon their privacy, despite having the capability to do so.

Third, the experts with lower levels of privacy concerns in particular mentioned that they gave little or no thought to the future consequences, subsequently deeming protective behavior superfluous. Group pressure and trust in others (e.g., the ‘if everyone does it should be fine’-heuristic; De Luca et al., 2016; Flender & Müller, 2012) eventually result in situations in which benefits outweigh risks. The experts explicitly stated that in many situations, the benefits are so attractive that the app is downloaded although their professional knowledge makes them aware of the potential risks. Although a lack of information and transparency deficit (e.g., due to the complex online environment and technical processes running in the background) are an important factor assumed to influence the behaviors of general users (Acquisti et al., 2016; Bräunlich et al., 2020), it is not surprising that these limitations are not considered by experts. Nonetheless, the results of our study confirm that even considerable technical knowledge is easily outweighed by internal and external factors. Similar to the behavior of general users, a lack of time, unwillingness to genuinely scrutinize an app and

trust in peers play important roles in influencing the decision-making processes of experts.

4.5.1.3 Determining the appropriateness of apps

The cues experts take into account for judging the suitability of apps differ from the strategies applied by general users. Similar to the findings of Jorgensen et al. (2015), requested permissions are by far the most frequently reviewed cues experts use, whereas general users may at best only superficially consider them (Felt et al., 2012). Experts see permission requests as the most informative cue with regard to potential privacy threats. However, the effects of their reviews of permissions are questionable, as most experts reported they still used apps that demanded questionable data privileges. Furthermore, even privacy and cybersecurity experts sometimes have a hard time understanding permission requests and assessing their necessity to app functionality requires considerable reflection. If experts have trouble using permission requests as useful cues, how can general users be expected to understand and use them?

One might expect that experts are inclined to make decisions based on their expertise and professional judgment, but the influence of others—in the form of reviews, downloads, and ratings—should not be underestimated. Other cues involve assessing app descriptions and the company or developer behind the app, which experts can probably categorize more easily than general users. Although their evaluation strategies may differ, experts are similar to general users when it comes to their ‘hit-or-miss’ analyses rather than an in-depth risk assessment. The knowledge they have of potential privacy risks and ways to avoid them plays no significant role here. To avoid ill-considered decision-making, with exposure to overlooked risks and unintended data disclosure, users need quick, comprehensible, intelligent, universal privacy-related information and warnings.

4.5.2 Practical implications

Our results suggest that the privacy knowledge experts possess plays little or no role in their privacy-related attitudes and behaviors. Some experts indicated that the broad and specialized knowledge they have is quite removed from the specific decisions they must make about downloading and using a specific app. Based on the results of our study, providing mobile phone users with more information about privacy and even making them more aware of the potential threats associated with downloading and using apps would not seem to be viable strategy when tackling the privacy paradox. General privacy knowledge and privacy awareness are relevant factors, but our expert study shows that even when both are present, other mechanisms will influence users to behave

in manners not in line with their privacy concerns. The privacy paradox must be seen as an unruly phenomenon, requiring more than information provision and persuasion for a solution.

In the short term, knowledge-based strategies such as privacy education and exhaustive privacy statements also appear to be of little help. A tool that might support users—both general and experts—in their privacy-related decisions would be a privacy visualization placed with all other app information in the app store. Ideally, such a visualization would quickly and comprehensively address a users' privacy issues just before app acquisition. It should inform users about the types of data gathered (e.g., personal or anonymous data) and the way the data are handled (e.g., analyzed, sold, protected, or stored) in a concise and transparent manner. The goal would be to highlight all relevant information without overwhelming users with too much information.

4.5.3 Limitations and future work

This study gained insights into the privacy considerations and online behaviors of users with expertise in privacy and cybersecurity. In interpreting the results, it is important to keep the following three limitations in mind.

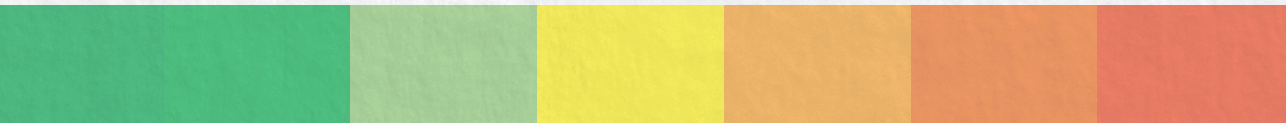
First, we included experts based on their expertise related to their professional roles. Although we were diligent in selecting participants—paying attention to their educations, their jobs and their experience—their specific levels and types of expertise as they pertain to privacy protection were not empirically ascertained. We can assume that the knowledge and views of our participants may be based on different professional and personal experiences, which we could not include in the interviews. Future research addressing the specific knowledge and views on privacy as it pertains to experts could for instance, take place in a Delphi study and prove an interesting follow-up to our study.

Second, retrospective face-to-face interviews are by their very nature vulnerable to social responsibility and memory bias. As a result, the behaviors reported by participants might deviate from their actual behaviors. We tried to reduce social desirability bias by creating a confidential atmosphere in the interviews and—given the results of the interviews—are confident that we succeeded in this: Two-thirds of the participants admitted to paying only moderate to very limited attention to their personal online privacy, not inflating their valuation of privacy. We tried to reduce memory bias by encouraging the participants to come up with specific examples. Future research might try to overcome these potential sources of bias by observing the actual behaviors of users, within real world or controlled scenarios.

Third, we compared the experts' knowledge and self-reported behaviors only to the aggregated insights from earlier research, choosing not to conduct the same study among experts and general users. It would be interesting if future research would make a direct comparison between the privacy behaviors of cybersecurity experts and general users. This can be done in a qualitative study based on things such as scenario-based interviews, observations, or in a quantitative pseudo-experimental design.

4.6 Conclusion

The purpose of this study was to examine the privacy perceptions and online behaviors of privacy and cybersecurity experts. The results show that the privacy evaluation and reported online behavior of experts is comparable to that of general users. Despite their technical backgrounds and thorough understanding of privacy risks, the majority of experts seldomly behave in more precautionary manners online. Instead, experts often engage in 'hit-or-miss' app analyses, making them just as vulnerable to heuristic thinking, immediate gratifications, or optimistic bias as their lay counterparts. Our results suggest that a heightened awareness of privacy threats and potential data abuse does not significantly diminish the potency of the privacy paradox among users.



5

Toward an understanding of online privacy perceptions: Using the Q-sort method to identify different user perspectives

Barth, S., Ngo, T., De Jong, M. D. T., & Krämer, N. C. (2020). *Toward an understanding of online privacy perceptions: Using the Q-sort method to identify different user perspectives*. Manuscript submitted for publication.

5.1 Introduction

Smartphones have become a mainstay technology in everyday life. Users are accustomed to 24/7 connectivity, button-press purchasing, and the practicality of many mobile applications (referred to as apps from here onwards). However, the usage of apps often requires disclosure of personal data. Research shows that users have concerns about their online privacy but tend to disclose personal information seemingly without hesitation (Barth et al., 2019; H. J. Smith et al., 2011), a phenomenon known as the privacy paradox (Barnes, 2006). Research has predominantly focused on generic explanations for this paradoxical behavior, assuming that (1) users more or less rationally weigh the costs and benefits of data disclosure (and the perceived benefits often outweigh their privacy concerns), (2) users often cannot adequately estimate privacy risks due to a lack of knowledge or shortcomings in their decision-making process, or (3) users may not even take the privacy aspects of their online behavior into consideration (Barth & De Jong, 2017). The available research largely neglects the individual and context-dependent aspects of online privacy perceptions, even though it is plausible that not all kinds of personal data are seen as equally sensitive and users may differ in their judgments of privacy threats. These assumptions are in line with the theory of contextual integrity (Nissenbaum, 2004, 2011), which assumes that users' perceptions of online privacy may differ from one context to another. What users tend to disclose in one particular situation may not apply under different circumstances. Recent studies on the context dependency of privacy perceptions mainly focused on social networks and other online media but less is known about the mobile computing context.

The research reported in this study aims to provide a better understanding of users' privacy perceptions in a mobile environment. We begin by assessing whether users' privacy perceptions are affected by the type of app involved (comparing a health and a news app). We then investigate whether it is possible to identify groups of app users who have similar views on privacy, focusing on three contextual factors: the types of personal information collected (*what*), the way the information is processed (*how*), and the party using the information (*who*). We did this using the Q-sort method, a research approach developed for segmenting groups of people based on the structures of their views (S. R. Brown, 1986, 1993; Stephenson, 1953).

5.2 Earlier research

The emergence of mobile computing has drastically changed our understanding of information privacy. Static information sharing developed into a fluent exchange of personal information (Barkhuus, 2012; Nissenbaum, 2004). Properties of smartphones like location tracking, 24/7 connectivity, and users' perception of their smartphone as an extension of their body (Shklovski et al., 2014) create personal data that were previously unthinkable and unsharable. As users have gotten accustomed to the affordances of mobile apps and have a basic understanding that some of these affordances may require some degree of privacy threats, universal privacy norms are no longer feasible, as users might judge the potential privacy violations of certain apps differently than those of others and different users might do so in different ways. This makes the operationalization of online privacy very hard as it seems to be context-dependent and determined on the individual level: For instance, while, for healthcare apps users might be inclined to disclose sensitive and intimate personal data, such as symptoms or heart rate, to get health-related insights, they might not want to disclose this information in the context of news apps.

While such privacy perceptions can be framed in terms of costs and benefits within the privacy calculus theory (Dinev & Hart, 2006), the validity of this approach is limited by information asymmetry (K. Martin, 2013). Users are not able to engage in a mere cost-benefit-analysis as (1) they often do not have all necessary information at their disposal, and (2) their information processing capacity is bounded by nature (Simon, 1972). Context dependency and bounded rationality raise the questions of what information is important to users regarding their online privacy within specific contexts, and how privacy issues can be communicated without overwhelming users with too much information. Taking a normative and universal approach and solely highlighting potential risks to one's online privacy might not be sufficient. A user-centered approach for operationalizing privacy is needed to adapt to users' privacy preferences and to avoid denial of such information like in privacy statements (Beldad et al., 2010). To overcome the limitations of the privacy calculus, we take another approach which frames privacy as contextual integrity.

5.2.1 Privacy as contextual integrity

According to Nissenbaum (2011), the roots of online activities are highly integrated in one's social life in general. Individual privacy boundaries and perceived norms that are shaped in offline social contexts, interactions and social roles are expected to be translated into the online context as well.

However, the properties of the online environment exceed what is possible in offline settings and mobile computing seems to create even more possibilities. Capturing implicit and explicit disclosed personal information, the analysis and dissemination thereof blurs the boundaries between privacy norms. One would expect that similar to the offline environment, online footprints should not be collected, registered, or analyzed. However, recent developments show that privacy laws associated with the offline context are seldom translated into the online environment. Still, the marketplace and commercial context is taken as the proxy for online privacy operationalization and protection, leaving many facets of the online context apart. It seems that data is the new currency while the product or service the user is aiming for just fades into the background (K. Martin, 2016b). Eventually this leads to a situation where the user leaves behind his or her privacy preferences, which leads to anxiety, uncertainty, feelings of helplessness, and fear (Nissenbaum, 2011).

According to the theory of contextual integrity (Nissenbaum, 2004, 2011), every aspect of life involves certain norms of information flow. Some norms are explicitly defined and others are implicit, fluent or even not fully developed. Furthermore, such norms must be defined in terms of appropriateness (what information is appropriate to disclose) and information distribution (how is information distributed and does this information flow respect contextual norms). Informational norms and, therefore, privacy expectations might differ from one context to another and are variable over time. Hence, what is seemingly appropriate to disclose in one context, might not meet users' privacy expectations in another one. Users seem to strive toward maintaining their privacy boundaries and norms according to the context. Contextual integrity is violated if either the norm of appropriateness or the norm of distribution or both is breached. As with attitudes of people in general, privacy perception is not entirely stable and might change in the course of time, influenced by technological developments, habits, or learning effects, for instance. Furthermore, the information in question, how it connects to the context, the actors involved in the information exchange and their relationships to each other, pre-defined rules of information distribution and changes to that distribution over time, shape the context and determine contextual integrity of information disclosure (Nissenbaum, 2004, 2019).

5.2.2 Privacy as a social contract

In her work on the social contract approach to privacy, K. Martin (2016a) added a moral layer to the view of contextual integrity (Nissenbaum, 2011). Rather than focusing on actual negotiated privacy norms, unstated agreements between parties involved in information exchange are made that are based on moral

values. Here, individuals share information within a particular community and in accordance with norms that regulate information sharing and disclosure. However, this view goes beyond gathering consent from the user but shifts information exchange toward mutual agreement and maintenance of beneficial relationships between all stakeholders involved in the information exchange (K. Martin, 2016a).

In the social contract approach to privacy, information sharing online is no longer an all-or-nothing approach of either you share information and lose the right to privacy (access view) or permit information disclosure and maintain privacy (control view). Privacy as a social contract allows users to not only (re-)gain power about with whom to share what information and for what purpose but also allows them to differentiate between relationships and adjust their privacy preferences accordingly. Furthermore, discriminately sharing of information allows users to disclose information to others without relinquishing privacy but maintaining privacy expectations around disclosed information (e.g., how that information is used, for what purpose and by whom; K. Martin, 2016a; K. Martin & Shilton, 2016). Based on contextual integrity, privacy as social contract distinguishes between four key parameters for informational norms, also known as contextual factors:

- *Why*: use cases; the context or circumstances of the application
- *What*: the types of information that are transmitted
- *How*: transmission principles; the constraints on the flow of information with regard to distribution, dissemination, and transmission
- *Who*: actors; the senders and recipients of information, and information subjects, could be single or multiple individuals, or organizations

To this date, only a few empirical studies have applied the theory of contextual integrity to investigate specific contexts. K. Martin and Shilton (2016) used vignettes to identify privacy expectations in the context of mobile devices. They found that many of the common data processing procedures used by mobile app companies do not meet users' privacy expectations. Furthermore, they demonstrated the importance of contextual factors as they pertain to user privacy expectations. Shi et al. (2013) used contextual integrity to examine interpersonal information boundaries on social network sites. They found that violations of the contextual factors can result in privacy concerns, highlighting the importance of the alignment of contextual factors for privacy perception. Using the theory of contextual integrity, Apthorpe et al. (2018) conducted a study on privacy norms

and derived design recommendations for IoT device manufacturers, emphasizing the usefulness of contextual integrity when identifying privacy norms. Hence, by applying the view of contextual integrity, it is possible to determine both specific and implied results as they pertain to different contexts for privacy.

In addition to the theory of contextual integrity, individual characteristics such as personality traits (agreeableness, openness to experiences, and conscientiousness), self-efficacy, and risk-taking might influence privacy needs in terms of perceived privacy violations (Junglas et al., 2008; Korzaan et al., 2009; Osatuyi, 2015; Taddicken, 2014). As it is impossible to consider the privacy needs of all individuals, we aim to find user groups that can be generalized for the use cases (*why*) and corresponding contextual factors (*what, how, who*). While a quantitative survey is an effective and widely used method to assess privacy norms and expectations (Shvartzshnaider et al., 2016), surveys aim at generalizing assumptions within a given population. However, under the premise that privacy is highly context-dependent, contextual factors might change within a given use case and might be influenced by individual characteristics. Furthermore, applying a generalizing research method seems insufficient for privacy visualizations, while entirely relying on individual characteristics is also problematic. Hence, we deem the identification of user groups as the most appropriate way to obtain a balance between generalization and individual user characteristics. To achieve this, we used a novel method in privacy research, the Q-sort method. This method allows to make result-based assumptions on privacy perceptions within a specific use case as it takes contextual factors and individual characteristics into account. Consequently, the research questions are:

RQ1. To what extent do contextual factors (*what, how, who*) differ between use cases (*why*; here: health vs news)?

RQ2. Which user groups, derived from individual characteristics, correspond to the contextual factors (*what, how, who*) and use cases (*why*; here: health vs news)?

5.3 Method

To investigate the research questions, we conducted a series of Q-sorts. A Q-sort is a session in which participants individually rank a set of items in a quasi-normally distributed pyramid-shaped matrix (S. R. Brown, 1993; Stephenson, 1953). The analysis typically involves a factor analysis in which participants instead of items are correlated. Our study consisted of separate Q-sorts for

three contextual factors (*what*, *how*, and *who*). The fourth factor (*why*) was manipulated experimentally: Half of the participants were placed in the context of a health app, the other half in that of a news app. In the analysis, we first focused on differences and similarities between the health and the news app. After that we tried to identify user groups with similar views on privacy for the contextual factors *what*, *how*, and *who*. The study was approved by the ethical committee of the department of Computer Science and Applied Cognitive Science, faculty of Engineering, University of Duisburg-Essen.

5.3.1 Between-subjects variable: app type

To investigate the effect of app type (the *why* contextual factor) on online privacy perceptions, we experimentally manipulated the context in which participants completed the Q-sorts. Half of the participants thought of privacy in the context of a health app (a pedometer); the other half of them were placed in the context of a news app. Both apps aim at tailoring services based on personal information, but differ in the types of data they collect and the way they collect them. The health app records intentionally provided generic personal data to monitor exercise patterns throughout the day; the news app utilizes unintentionally provided specific data such as search results, entered keywords, or click histories, to provide tailored news content. The use of unintentional data puts users in a powerless position and makes them highly dependent on the system that processes personal data (Pangrazio & Selwyn, 2019). Due to these differences in app characteristics, we expected that there might be differences in users' privacy perceptions. Participants received a health-app or news-app scenario for the three sorting tasks (see Appendix 5.1 for the instructions). The Q-sort materials were exactly the same for both conditions.

5.3.2 Research materials: Q-sort method and questionnaire

Three separate Q-sets were developed, focusing on the three contextual factors (*what*, *how*, and *who*). The items (Q-sets) were formulated as privacy-related actions, which participants had to place in a matrix from “no violation” (-3) to “very strong violation” (+3) (see Figure 5.1 for the Q-grid used in all three Q-sorts).

To develop the Q-sets, two steps were taken. First, all possible action statements for the three contextual factors were gathered (the *concourse*). Items were derived from Nissenbaum's (2011) description of each contextual factor and from statements formulated by K. Martin and Shilton (2016). Items were formulated to fit equally well in both conditions.

We extended the existing statements of the contextual factors of *what* and *how*. While K. Martin and Shilton (2016) distinguished between using information for future ads, selling data to an online auction, and using information for social advertising, we also added aspects of encryption, identification, permissions, and storage in our study. In addition, we diversified the statements of the type of data (*what*). Secondly, from the concourse, three subsets of statements were selected, consisting of 16 statements each, which were eventually presented to the participants (the Q-sets; see Appendix 5.2).

Additionally, a short questionnaire was made focusing on participants' background characteristics. This was used to explore whether there were noteworthy patterns of characteristics in the user groups identified by the Q-sorts. The questionnaire focused on:

- Age, gender, nationality, education, and educational background (five items);
- Duration of smartphone ownership (one item; eight-point scale going from 1=less than a year to 8=more than 7 years);
- Past behavior regarding denial of permissions (one item; five-point scale going from 1=never to 5=always, plus the additional answer “I am unaware of that option”);
- Specific privacy concerns: downloading mobile apps (one yes/no item plus the additional option “I don't know”).

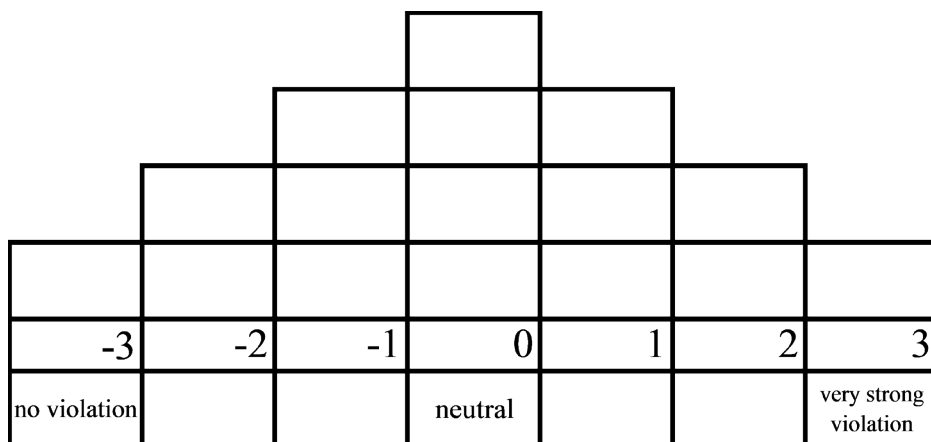


Figure 5.1 Q-grid containing 16 fields and going from “no violation” (-3) to “very strong violation” (+3)

5.3.3 Procedure

Data were collected in individual face-to-face sessions with participants. At the start of a session, participants were randomly assigned to one of the two conditions (health vs news). After reading the study information and signing the consent form, participants were briefed on the procedure. With their specific scenario (health or news) in mind, participants then completed the three Q-sorts. The order of the three Q-sorting tasks was randomized. After that, they were asked to complete the questionnaire. The sessions ended with a debriefing.

5.3.4 Participants

In total, 100 German participants were recruited: 50 for each condition (health vs news). The convenience sample was recruited through online social networks, forums, and via a university participant pool. Of the 100 participants, 27 were male and 73 female. Their ages ranged from 18 to 72 years ($M = 29.5$, $SD = 13.0$). Of the sample, 52% held a high school diploma, 19% a professional degree, and 20% a bachelor's degree or higher. Most participants had owned a smartphone for more than six years. About half of the participants (52%) indicated having privacy concerns when downloading mobile apps on their smartphones; the other half of the sample was either not concerned (27%) or did not have an opinion (21%). As far as app permissions are concerned, less than 50% of the participants claimed to always or predominantly deny them.

5.3.5 Analysis

5.3.5.1 Comparison of health and news app

To test for differences between the health app ($n = 50$) and news app ($n = 50$), we conducted t-tests for independent samples with all statements of the three Q-sets, which amounted to a total of 48 unpaired comparisons. In addition, Q-sort factor analyses with varimax rotation were conducted for each app type separately, using PCQ for Windows (version 1.4). These analyses were used to decide whether the data of both apps would be merged or not for the remaining analyses. As will be described in the results section, our analyses led to the conclusion that the data of both apps would be merged.

5.3.5.2 Identification of user groups

The Q-sort data regarding three contextual factors (*what*, *how*, and *who*) were analyzed separately using PCQ for Windows (version 1.4). In each analysis, a correlation matrix of all Q-sorts was made, representing the level of (dis) agreement between individuals, and a centroid factor analysis with varimax rotation was conducted to identify similar groups.

In all three factor analyses, a substantial number of participants did not load significantly on one of the factors (40, 42, and 43 participants, respectively, for the *what*, *how*, and *who* factor). In addition, three participants in the analysis of the *who* factor had significant loadings on more than one factor (confounded participants). All these participants were omitted from further analysis. While this amount of exclusions might seem drastic, previous research has shown that stabilization in the Q-sort method is usually reached at a threshold of approximately ten participants that load significantly on at least one factor. Consequently, it is possible that a considerable number of Q-sorts are idiosyncratic and do not load on any factor (Fairweather 2001). Eventually, more than 50 participants were included in all three Q-sort analyses, which can be considered a sufficient sample size (S. R. Brown, 1986). The number of participants omitted is not unusual and does not threaten the validity of the results.

Based on the calculation of the correlation matrix of all Q-sets exploratory factor analyses with varimax rotation were performed, starting with nine factors and gradually reducing the number of factors. Eventually, three-factor solutions were chosen for all three contextual factors. All factors had eigenvalues higher than 1, accounting for 72% (*what*), 74% (*how*) and 64% (*who*) of the variance in the Q-sorts and all single factors accounted for more than 5% of the variance. After identification of three distinctive factors for each contextual factor, the groups were interpreted and described qualitatively. For each user group, a label was formulated that best represented its specific characteristics. The last step in the analysis involved a comparison of the user groups on the basis of the three contextual factors. Using a cross-tabulation analysis, we did not find any pattern in the way participants were categorized in the three contextual factors. In other words, participants who scored the same on the factor *what* did not score consistently the same on the factors *how* and *who*. We therefore concluded that the contextual factors *what*, *how* and *who* should be interpreted separately from each other.

5.4 Results

5.4.1 Comparison of health and news app

With only a few exceptions, the results of the t-tests showed no statistical differences between the two app types with regard to perceived violations of online privacy. Analyzing all three contextual factors, we only found three statements that differed significantly between the app types. For the factor *what*, the item ‘disclosure of photos’ was considered to be more violating for the news app ($M = 2.10, SD = .84$) than for the health app ($M = 1.60, SD = 1.14$) ($t(98) = -2.49, p < .05$). A Cohen’s d of .50 indicates a medium effect size. For the factor *how*, the item ‘personal data are transmitted with personal identifiers’ was rated as more violating for the news app ($M = 1.28, SD = 1.33$) than for the health app ($M = .66, SD = 1.44$) ($t = -2.24, p < .05$). A Cohen’s d of .45 indicated a medium effect size. For the factor *who*, the item ‘my personal data are analyzed by the operating system (e.g., iOS by Apple, Android by Google)’ was seen as more violating for the health app ($M = -.06, SD = 1.11$) than for the news app ($M = -.70, SD = 1.20$) ($t = 2.67, p < .05$). A Cohen’s d of .55, again, indicates a medium effect. The remaining 45 comparisons were nonsignificant. Furthermore, we compared the outcomes of separate Q-sort analyses for each contextual factor of the health and the news app. These analyses did not result in differences in the nature of the groups identified. We therefore conclude that the factor *why*, operationalized by our distinction between a health app and a news app, did not significantly affect participants’ privacy perceptions. As a result, we merged the data for the two apps for the analyses described below.

5.4.2 User groups based on types of information (*what*)

Regarding types of potentially privacy-sensitive information, three user groups were found with apparently different perspectives on online privacy. The groups seem to largely agree on unacceptable practices, such as access to their photos, home address, or personal messages, but differ in practices that would be more acceptable to them.

Group 1: Disclosure of anonymous data is acceptable. The first group predominantly focuses on the anonymity of the data disclosed ($N = 30$; mean age = 29.3; composite reliability = .99; S.E. of factor Z-scores = .14, eigenvalue = 29.7 with a variance of 30). Users in this group tend to accept disclosure of anonymous data, but see information that might identify individuals as a strong online privacy violation. The distribution of statements distinctive for this group shows a rising line from non-identifiable data to data that enable definite identification. On the one end of the spectrum, anonymous demographical

data such as gender, age, and height are not perceived as a privacy violation. Data linked to online behavior, such as history of online purchases, search results, and numbers of clicks are considered as moderate privacy violations. On the other end of the spectrum, information that could lead to the identification of specific individuals such as home address, photos, or messages were perceived as severe privacy violations. The background variables of users in this group suggest that many are unsure of how to deal with the topic of online privacy. One third of the participants are uncertain whether they have privacy concerns. The majority did not use protective mechanisms, such as denying permissions to apps, despite having substantial experience with smartphones (> 6 years).

Group 2: Tracking of online behavior is acceptable. The second group acknowledges the fact that tracking online behavior is almost unavoidable in online services such as mobile apps ($N = 19$; mean age = 26.9; composite reliability = .98; S.E. of factor Z-scores = .18, eigenvalue = 23.0 with a variance of 26.9). These users seem to be familiar with online processes and know that certain data about online behaviors are always gathered. Therefore, they rate number of clicks, history of online purchases, and search results as less violating to their online privacy. Non-specific data such as gender, height, and age were considered to be acceptable as well, whereas potential identifiers with moderate informative value about a person, such as birth place, first and last name, contact list, or mobile phone number, were placed in the middle of perceived online privacy violations. However, information that can be considered as unique identifiers, such as home address, messages, or personal photos, was perceived as a major privacy threat. Thus, certain things are considered to be business as usual online: What you need to disclose in order to use a certain app is acceptable as long as it is not identifiable and cannot be linked to a specific person. Users in this group are not only aware of the fact that some types of information are collected and processed, but accept it that a given amount of personal data is the normal 'price' for usage. The background variables of the users in this group suggest that they are relatively savvy when it comes to online apps and online privacy. They have sufficient experience with using smartphones (> 6 years) and consciously consider privacy aspects of apps, only downloading apps when they meet their cost-to-benefit expectations. Users in this group take a proactive approach in controlling the flow of their personal data and mostly deny app permissions.

Group 3: Equating online and offline privacy. The third group equates their online privacy boundaries with the ones they would have in their offline lives ($N = 19$; mean age = 26.1; composite reliability = .97 with S.E. of factor Z-scores = .23; eigenvalue = 18.7 with variance of 19). Identifiers that are

easily given away in real-life encounters, such as first and last name or gender, are deemed acceptable. Information that requires some level of trust to share with others, such as weight, home address, contact list, and mobile phone number, are considered to be more sensitive. Information that people normally only share with others with specific purposes in mind, such as photos, real-time location, or messages, are considered as very private and, if unintentionally disclosed, as strong privacy violations. Typical data involved in tracking online behavior, such as number of clicks, search results, or online purchases, are seen as privacy threats as well. The background variables of the users in this group suggest that they particularly value the control of their personal data. The majority of the users indicate that they often deny permissions to apps. However, when asked directly, they generally expressed to not have privacy concerns.

5.4.3 User groups based on the way information is processed (*how*)

Regarding the way personal information is processed, three user groups were found with different perspectives on online privacy. Again, the user groups seem to agree on practices that they consider to be severe privacy violations, such as selling or sharing their data with third parties, but differ in their views on practices that are more acceptable to them.

Group 1: Pro personalization. The first group is willing to disclose certain personal information supporting online services that are tailored to their needs and preferences ($N = 18$; mean age = 29.4; composite reliability = .98 with S.E. of factor Z-scores = .18; eigenvalue = 24.2 with variance = 24). These users consider the transmission of personal data without their permission, the usage of personal data for purposes other than declared, and the transmission of data including personal identifiers as a strong privacy violation. However, such practices are seen as less violating if personal data are transmitted with permission in an encrypted form for purposes such as targeted advertising. It seems that despite privacy concerns and a desire to retain control over their personal data, they appreciate the advantages of personalized online services such as targeted advertising. They accept that responsible personal data disclosure is a calculated and strategic necessity when using online services. Thus, they want to know where their data are stored and for which purpose they are used. Staying in control is an important requirement for doing things online. The background variables of the users in this group shows that the majority has privacy concerns. However, their actual behavior regarding the denial of permissions to apps underlines their ambivalence in the privacy-benefits trade-off: Some indicate to deny permissions irregularly, others range from rather seldom to half of the times.

Group 2: Against commercial data use. The second group wants to retain control of their personal data and is against the use of their data for commercial purposes ($N = 29$; mean age = 28.1; composite reliability = .99 with S.E. of factor Z-scores = .14; eigenvalue = 27.6 with variance of 28). These users consider local data storage, data storage within a limited timeframe, and the disclosure of personal data with permission as minor privacy violations. Targeted advertising, profiling, and transmission without security measures are placed in the middle range of perceived privacy violations. However, the transmission of personal data without permission, the use of data for purposes other than declared, and particularly the selling of data to third parties are seen as strong online privacy violations. Shady data-related practices beyond the user and the service-provider and making profit from their personal data are no-go's for them. The background variables of this group of users appear to be mixed. About half of the participants indicate having online privacy concerns and the behavior regarding permissions ranges from seldom to always denying.

Group 3: Against identification. The third group does not accept data aggregation from different sources that might lead to personalization of services and therefore to the identification of users ($N = 11$; mean age = 29.1; composite reliability = .97 with S.E. of factor Z-scores = .23; eigenvalue = 21.6 with variance of 22). These users want to be in control of their personal data, placing much emphasis on local data storage and transmission only with permission. Data aggregation from different sources and commercialization that is linked to identification and personalization are unacceptable for this user group. Data used for purposes other than declared and without permissions are perceived as strong violations of online privacy, whereas data that can be traced back to a specific user are perceived as the worst breach of their privacy boundaries. The background variables of this group of users was balanced. Roughly half of them had privacy concerns and frequently denied app permissions.

5.4.4 User groups based on the parties with access to their information (*who*)

Although our analysis revealed different user groups, a closer look at the group characteristics showed that users focus on the actual handling of their personal data rather than on the party responsible for it. All user groups see the dissemination, including the selling, of personal data as most threatening to their privacy, whereas the clusters of statements involving access, gathering and analysis of personal data varied along the continuum of perceived privacy violations (going from non-violating to extremely violating). In other words, users pay considerably less attention to the parties involved than to the practices

themselves. It is imaginable that participants are unable to differentiate between parties involved and therefore perceive the practice itself as more salient than the agents involved. Eventually, we decided to not further interpret the groups corresponding with the factor *who* as the practices themselves—how data are actually treated—are already covered by the contextual factor *how*. The lack of attention users have for the agents involved in the practices, however, is an interesting finding in itself.

5.5 Discussion

5.5.1 Main findings

Our comparison of types of app (*why*; health vs. news app) and our analysis of the contextual factors *what*, *how*, and *who* resulted in four insights that contribute to our understanding of users' online privacy perceptions (see Figure 5.2). First, the type of app involved, operationalized as the distinction between a health and a news app, does not appear to play an important role in users' privacy perceptions. Our research suggests that users consider online privacy from largely the same perspectives in both scenarios. Of course it should be noted that we only compared two app types and that it is imaginable that privacy considerations become more salient and therefore different in the case of apps that are inherently more privacy-sensitive (e.g., health apps beyond physical exercises, news apps in nondemocratic environments, financial apps).

Second, in their privacy considerations, users do not seem to pay much attention to the agents involved in privacy-sensitive practices (*who*). They primarily focus on the other two contextual factors: which personal data are collected (*what*) and what is done with them (*how*). Our Q-sort regarding the factor *who* did not result in a meaningful segmentation. Contrary to our expectations, participants ignored the different parties included in the Q-sort and instead focused on the privacy-related practices themselves. This suggests that they are either unaware of the possible agents involved or do not care about them.

Third, the segmentation of user groups with differential views on privacy can be attributed to different views on practices that are considered to be more or less acceptable, not to different views on what is not acceptable. In all Q-sorts, participants largely agreed on practices that they saw as strong violations of online privacy. For the factor *what*, for instance, access to their photos, home address, and messages were no-go areas for all users. For the factor *how*, the

aspect of betrayal played a prominent role: Selling data or sharing them with third parties were considered to be the strongest violations of online privacy.

Fourth, the Q-sorts regarding the types of information (*what*) and the way the personal information is processed (*how*) resulted in the identification of three different user groups each. Regarding the factor *what*, the three groups had the following starting-points: (1) anonymous data are acceptable, (2) tracking online behavior is acceptable, and (3) online privacy needs are just like offline privacy needs. Regarding the factor *how*, their starting-points were: (1) personalization is acceptable, (2) commercial data use is unacceptable, and (3) personal identification is unacceptable. These viewpoints represent different perspectives on online privacy that are easily underexposed in survey and experimental research on online privacy.

5.5.2 Theoretical implications

Our findings can be seen as a partial confirmation of the contextual integrity view on online privacy. They show that users differentiate in their views on the seriousness of online privacy threats and that different groups of users do so in different ways. We call it a partial confirmation, because (1) two of the contextual factors (*why* and *who*) did not seem to affect users' privacy perceptions, and (2) differentiations only involved privacy-sensitive practices that users consider to be acceptable, not practices that are seen as strong privacy violations.

To elaborate on the first side note: Our results show that not all contextual factors are equally prominent. The type of information (*what*) and the way information is processed (*how*) appear to be salient contextual factors for mobile phone users. Users' perceptions of privacy violations are affected by these two user-centric factors. Our finding that the type of app (*why*) does not matter is surprising, contradicting earlier findings by K. Martin and Shilton (2016) which suggested that such contexts do make a difference. More research on app characteristics that matter for users' privacy perceptions is needed. Our finding that the parties with access to personal data (*who*) do not matter in users' privacy perceptions is in line with K. Martin and Shilton (2016), who also found that agents involved only play a tangential role. Acquisti et al. (2016) and Bräunlich et al. (2020) noted that many users are unaware of privacy-related practices in online environments. As suggested by the contextual factors *what* and *how*, users can reflect on and form opinions about types of personal data that may be collected and the way they are handled, but are not inclined to overthink the parties doing this. When researching or regulating online privacy, it therefore seems recommendable to mainly differentiate on the basis of these two contextual factors (*what* and *how*).

To elaborate on the second side note: Our results show that there is common ground among users regarding practices that are considered to be strong online privacy violations. For the contextual factor *what*, home address, personal pictures and text messages are the types of information with the greatest sensitivity. For the contextual factor *how*, users are particularly sensitive to feelings of being betrayed: Distribution of data to third parties without permission and transmission with personal identifiers and for purposes other than declared are seen as the most violating practices. The finding that users largely agree on which practices are strong privacy violations still confirms the contextual integrity view on online privacy in that some practices are more intrusive than others, but does not suggest that individual differences between users are meaningful. Individual differences only become meaningful when we look at the practices that users find more or less acceptable. Our research shows that these differences relate to coherent and comprehensive views on the nature of online privacy.

Our results suggest that the overall concept of online privacy, in the minds of users, is multifaceted, with a core of unacceptable practices and a periphery of practices for which different user groups have different levels of tolerance. Activities such as online search tracking seems to be acceptable for some users, while others draw the line at anything beyond anonymous demographical data. A representative study among German citizens on artificial intelligence in online environments also pointed in the direction of differentiation (Kozyreva et al., 2020). Users appreciate the benefits of personalized services such as entertainment, shopping or search engines, but dislike personalization when it comes to political advertising or news feeds. Even if the opinion about usage of sensitive, personal data for personalized advertising is found to be negative in general, revealing information such as age and gender is acceptable. However, disclosure of sensitive information such as religious, political and sexual orientation, information about personal traits or content of personal communication is deemed to be unacceptable, let alone if such information is used for personalization of online services. Our results show that it is possible to find a middle ground between overly generalized perceptions of online privacy and strictly individual views on privacy. The segmentations that emerged from our findings enable us to better understand the motivation users may have for specific views on acceptable practices.

Our results also have implications for discussions about the privacy paradox (Barnes, 2006). First, we found support for the all theoretical explanations of discrepancies between people's privacy concerns and their actual behavior (cf. Barth & De Jong, 2017). In support of the view that users try to weigh the benefits

of an app against the costs in terms of privacy threats, we could see various user groups accepting certain types of privacy-sensitive practices acknowledging that they are part of the deal. Similar to Kozyreva et al. (2020), the clearest example is the group accepting data collection and handling related to personalization of the online experience. Users are skeptical about personalization but at the same time appreciate services such as personalized online shopping or entertainment. In support of the view that users do not seriously consider privacy when deciding to download and use an app, we observed that participants found it hard to form an opinion about privacy. Many participants had difficulties ranking the statements in the Q-sorts, reporting that they had never given serious thought to such privacy aspects. When prompted to give the matter serious consideration, they saw most factors as potential violations of their online privacy. Many found participating in our research insightful and the statements thought-provoking.

A possible explanation of the privacy paradox that is currently underexposed is that the umbrella term online privacy may not have a consistent meaning to all users at all times. When asked about the overall importance of privacy, users may be inclined to focus on the possibilities of serious privacy infringements—the kinds of threats all user groups saw as unacceptable. When considering whether or not to download and use a particular app, they might predominantly think of minor privacy violations. To gain more insight in the nature and the underlying mechanisms of the privacy paradox it seems therefore important to dissect the generic online privacy concept and focus on the perceived likelihood and severity of particular privacy violations.

5.5.3 Practical implications: Design recommendations for privacy visualizations

Our research suggests that it is important to support users in their decision-making about downloading and using mobile apps. Many participants mentioned a lack of awareness of privacy aspects of mobile apps, indicating that they normally do not think of it, as well as a lack of knowledge. At the same time, they appeared to be willing and able to make sense of their online privacy when prompted by the research materials. Practical tools are needed to support users in awareness and knowledge acquisition about data handling practices such as data collection and data aggregation (Pangrazio & Selwyn, 2019). Insightful privacy labels or icons could be such a tool (Metzger, 2007). Our findings may be used to inform the design of such labels.

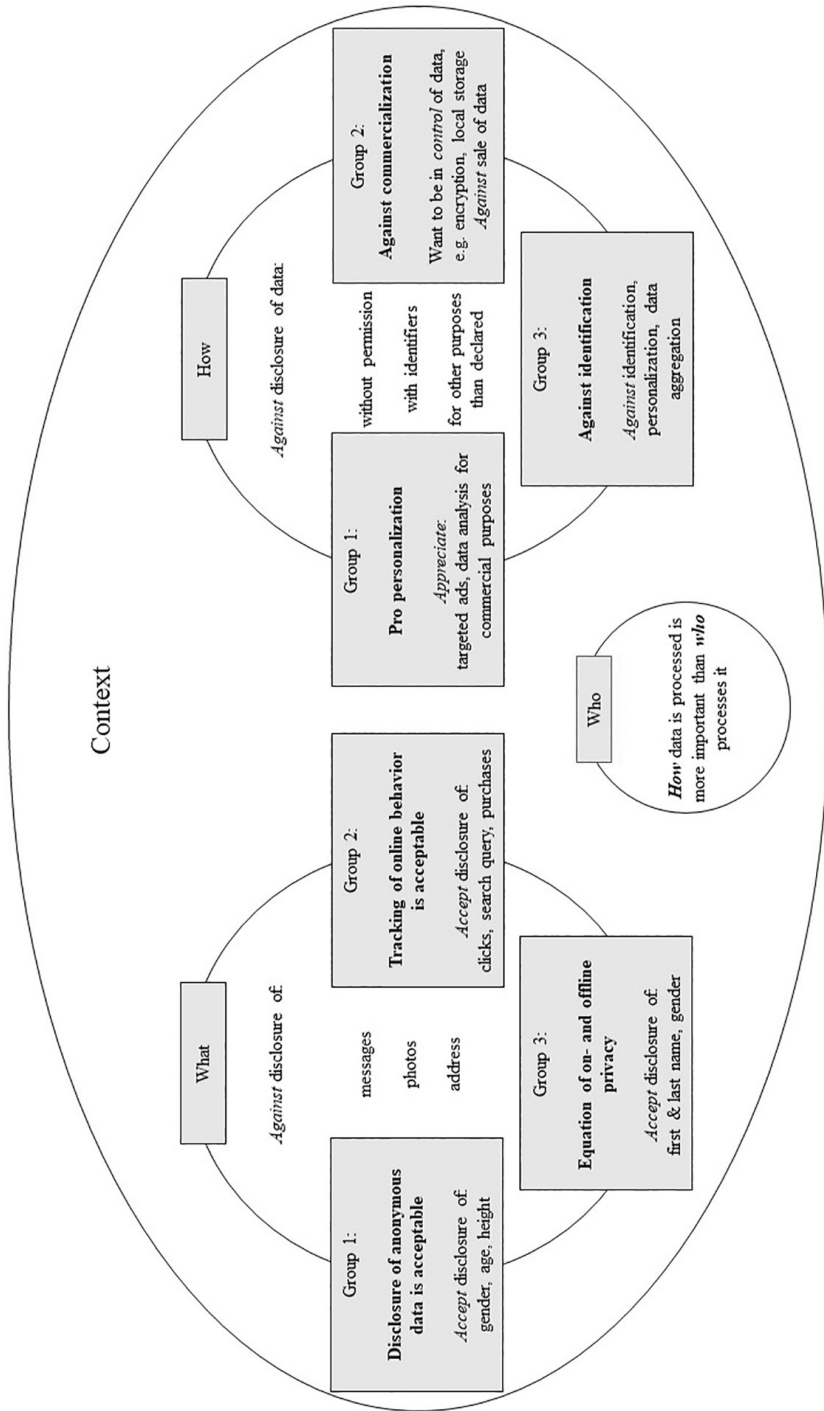


Figure 5.2 Visual representation of the results from Q-sort analysis

First, our research shows that it is possible to find a middle ground between ‘one size fits all’ privacy approaches and highly individual privacy perceptions. There is a core of practices that are broadly seen as unacceptable, and there are different but meaningful perspectives on practices that may be more or less acceptable. This calls for a flexible system in which users are prompted with an overall privacy rating, but also have the opportunity to explore the specific aspects that are potentially threatening their online privacy. The overall rating is relevant for creating awareness and supporting the decision-making of less-engaged users. The information on specific aspects is important to provide credibility and transparency to the overall rating and to serve users who have specific concerns about privacy or who are comfortable with the disclosure and use of certain types of personal data.

Second, the privacy label should particularly focus on the types of data collected (*what*) and the way this personal information is processed (*how*). The parties involved in the handling of personal data (*who*) can be ignored in the system. Of course, it should be emphasized that the handling of personal data comprises practices in which data are sold and shared with third parties. The selling or unauthorized sharing of data are considered to be unacceptable to users, but it is less important to users to whom the data are sold or with whom they are shared.

5.5.4 Limitations and suggestions for future research

There are some limitations that must be kept in mind when interpreting our findings. The first limitation involves the national context. The research was conducted in Germany and we cannot be sure that people’s perceptions are the same as those in other countries. International comparisons are needed to verify the generalizability of our findings. The second limitation involves the two types of apps used in our study. It would be interesting to investigate whether more extreme variations of app types would lead to differences in privacy perceptions, comparing on the one hand, apps that do not need privacy-sensitive data for their actual functioning (like many types of games) and on the other hand apps that collect information that by its very nature is privacy-sensitive (like dating, financial or health apps, that record more than just exercise patterns). The third limitation is that many participants did not load significantly on one of the factors distinguished. However, we included more than 50 participants in our Q-sorts, a sample size that is considered sufficient for Q-sort factor analysis (S. R. Brown, 1986). Moreover, stabilization in Q-sort factor analysis is usually reached after significant loadings of ten participants on at least one factor. Hence, it is not

unusual that a considerable amount of participants do not load on any factor (Fairweather, 2001).

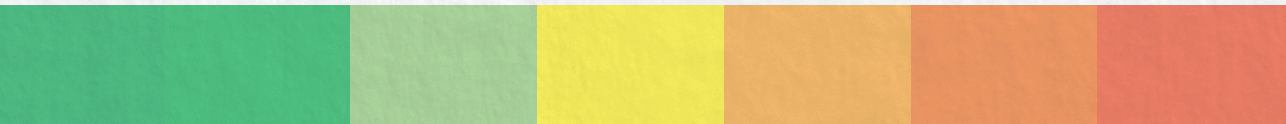
And our last limitation is that the Q-sort method forces participants to rank all statements in a matrix representing a normal distribution. Our results represent relative judgments about the severity of privacy violations, but cannot be seen as absolute judgments. Practices that end up at the ‘no violation’ end of the scale, may still be privacy threats to the participants.

In our view, future research should unravel the overall privacy concept into more specific threats. To gain a deeper understanding of the nature of the privacy paradox, we should know in more detail which aspects of privacy users value most and how they perceive the likelihood and severity of these specific privacy aspects. Starting point in such research would be that privacy perceptions may differ between different groups of users.

Another important line of research would focus on the effects of privacy labels on users’ perceptions and behaviors. To what extent are users inclined to pay attention to such labels, how do they make sense of them, and to what extent do they affect the privacy paradox?

5.6 Conclusion

In line with the theory of contextual integrity, our research found that two contextual factors—the types of data collected (*what*) and the way this information is processed (*how*)—play a significant role in users’ privacy perceptions. Users care about both factors and make differential choices in both of them. This led to the identification of in total six groups of users with different but coherent views on what could be acceptable in terms of their privacy-sensitive information. Two other contextual factors—the agents involved in the handling of privacy-sensitive information (*who*) and the type of app used (*why*) did not appear to be important. The differentiation of user groups was based on differences in practices that users find more or less acceptable, not on their views on the severe infringements of privacy. Users appeared to largely agree with each other on the strong violations of their online privacy. Based on our findings we argue that future research on the privacy paradox should focus more strongly on specific aspects of online privacy.



6

Understanding online privacy - A systematic review of privacy visualizations and Privacy by Design guidelines

Barth, S., Ionita, D., & Hartel, P. H. (2020). *Understanding online privacy – A systematic review of privacy visualizations and Privacy by Design guidelines*. Manuscript submitted for publication.

6.1 Introduction

Online services currently handle unprecedented amounts of user-related data (Schneier, 2015). Machine learning algorithms extract value from large amounts of data by recognizing hidden patterns, links, behaviors, trends, identities, and practical knowledge which has given birth to a 'big data economy' (A. L. Allen, 2016; Myers West, 2019). This has opened a 'Pandora's Box' of privacy concerns (Phelps et al., 2000; Tene & Polonetsky, 2011; H. Wang et al., 1998). But privacy policies are often lengthy, legally worded documents written to protect the provider (Antón et al., 2004; Fernback & Papacharissi, 2007). Even the interactive permission system found on modern smartphones fails to provide a sufficient understanding of the privacy risks involved with using an application (Benton et al., 2013; Y. Chen et al., 2019; Kelley et al., 2012).

To communicate privacy risks to users in a clear and concise manner, researchers, regulators, and industry have called for a more visual representation of how e-services handle personal data (Antón et al., 2007; Antón et al., 2004; L. Edwards & Abel, 2014; Holtz et al., 2011a; Rossi & Palmirani, 2017). Since 2001, the US Federal Trade Commission (FTC; Anthony, 2001) has been encouraging standardized, tabular privacy policies similar to nutrition labels. Carnegie Mellon's CyLab has developed and tested a 'privacy nutrition label' with promising results (Kelley et al., 2009). More recently, the European GDPR has suggested using 'standardized icons' to provide a meaningful overview of the intended data processing (The European Parliament and the Council of European Union, 2016). The Digital Advertising Alliance (DAA; 2016) displays a YourAdChoices button on their ads and the Entertainment Software Rating board has introduced icons indicating whether or not games share personal information with third parties (Haninger & Thompson, 2004). Researchers, Mozilla, and even the European Commission have proposed a variety of icons specifically designed to convey how personal data are handled. However, these visualizations differ with regard to the privacy attributes they cover, as well as their level of detail. Furthermore, the comprehensibility and effectiveness of the visualizations remains questionable as most of them have never been tested with users (Rossi & Palmirani, 2017).

Whereas visual representations of privacy attributes are intended for users, Privacy by Design (PbD) guidelines are intended for developers. Developers determine to a significant extent how user privacy is handled. Because developers are not privacy experts, they need clear and unambiguous instructions with regard to how personal data should be handled (Cavoukian, 2012). To begin with, developers need to know which privacy attributes are considered important by

users. While guidelines for what was once referred to as 'fair information practice' go as far back as the 1970s (Gellman, 2019), technological developments have prompted a renewed interest in developing privacy-aware information systems (Schaar, 2010). However, there is currently no generally accepted PbD standard or best practice. Rather, multiple regulators and industry stakeholders have each elaborated their own PbD principles which—similar to privacy visualizations—differ significantly in terms of the privacy attributes they consider.

So, *what are the most important, generally-applicable privacy attributes of online services?* We attempt to answer this question by systematizing knowledge surrounding privacy from relevant approaches in academia, industry, and government and by considering the opinions of both privacy experts and users, in order to compile, validate, and rank a complete list of privacy attributes. As a first step, a list of privacy attributes was derived by means of a systematic review of existing privacy visualizations and PbD principles. Second, this list was refined and extended in collaboration with information security professionals via interviews. Third, we distributed an online questionnaire among predominantly European privacy experts and users of online services, resulting in a ranking according to perceived importance from both perspectives. Finally, based on the results, we explain notable differences and patterns, and identify trends. Together, our results form a foundation for understanding, communicating, and discussing privacy, and inform the development of user-oriented privacy-aware online services. We present practical recommendations for the development of future privacy visualizations and PbD guidelines, as well as outline research challenges toward facilitating the analysis and comparison of privacy policies and investigating the context dependency of privacy attributes.

6.2 Background

The debate around privacy started in the late 19th century with the launch of the telephone and intensified throughout the 'cybernetic revolution' of the 70s (Miller, 1969). In his landmark 1967 book, Westin defined privacy as "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, 1967, p. 7). Fundamentally, modern privacy is about information (Cate, 1997). However, the concept kept expanding in both scope and significance with the emergence of the internet, mass surveillance, terrorism threats (H. Wang et al., 1998; Westin, 2003) and more recently with the development of big data and the Web 2.0 (A. L. Allen, 2016; Myers West, 2019; Phelps et al., 2000; Tene &

Polonetsky, 2011). Currently, privacy, and in particular online privacy, remains hard to define (Pavlou, 2011), or in the words of Solove, “a concept in disarray” (Solove, 2008, p. 1). H. J. Smith et al. (2011) notes that historically, privacy was seen as either a right, a commodity, a control, or a state. K. Martin (2016a, p. 557) sees privacy as a bilateral contract about “what, to whom, and for what purpose” information is gathered or disclosed within a given community and context. Nissenbaum (2004, 2011) posits that privacy is shaped by social boundaries and norms and—because individuals cannot provide truly informed consent—she suggests articulating context-specific norms which govern the collection and sharing of data online. According to her theory of contextual integrity, whether or not an action constitutes a violation of information privacy depends on variables related to the context, the nature of the information, the actors involved and their relationships to the data subject, as well as the terms for collecting and sharing information. Acquisti et al. (2016) discuss the economic value of privacy and also find that in some situations data sharing can be beneficial for the user, while in other situations it can be damaging. Nevertheless, in his landmark articles, Solove (2002, 2008) points out that while it is not feasible to arrive at an overarching definition of privacy, the concept can be understood by isolating common ‘essential’ or ‘core’ characteristics.

Solove (2008) approached this from a legal perspective by developing a taxonomy of privacy violations pertaining to information collection, information processing, information dissemination, or invasion. From a technical perspective, privacy metrics are often used to compute the efficacy of privacy-enhancing technologies (Wagner & Eckhoff, 2018), but these are of little use to people without a background in statistics. Morel and Pardo (2020) examined privacy policies and identified three dimensions of expressing privacy: natural language (from law), graphical (from organizations and privacy advocates), and machine-readable (from academia). However, they do not discuss which aspects of privacy should be covered by each dimension. Anwar et al. (2018) found commonalities between privacy laws and standards, but noted differences in nature and scope which require further investigation. Martín et al. (2014) highlighted a lack of technical privacy requirements and criticized disagreement between high-level privacy principles. According to Morales-Trujillo et al. (2018), in order to address privacy during software development and to be able to respond to user’s privacy concerns, a conceptual framework is needed that goes beyond data minimization and access control.

Acquisti et al. (2017) saw potential in efforts toward assisting users with online privacy decisions by helping them reflect on their actions before the fact or by ‘nudging’ them toward certain behaviors. But Rossi and Palmirani

(2017) concluded that existing privacy visualizations vary in terms of the privacy attributes they cover and criticized that the majority were not user-tested. They suggest a visual layer summarizing the privacy policy with special focus on the privacy principles of transparency and informed consent but to date, no new system has been developed. Motti and Caine (2016) reviewed icons related to privacy and classified them as either data collection, data transmission, data storage, data sharing, or access control.

Overall, there appears to be a lack of agreement in terms of decomposing privacy into its core attributes. In order to help understand online privacy, we identified a list of unified privacy attributes and ranked this list based on importance. We did so by systematically comparing proposals for conceptualizing privacy aimed at users (privacy visualizations) and at developers (PbD guidelines), considering all sources (academia, industry, and government), and accounting for the perspectives of users as well as of privacy experts.

6.3 Method

The aim of this study is to distill, validate, and rank a complete list of privacy attributes. The first step toward achieving this was to perform a systematic review to identify privacy visualizations (Section 6.4.1) and PbD principles (Section 6.4.2) relevant for online services. We then extracted a list of privacy attributes by coding the results until reaching satisfactory inter-coder reliability and then refining it with practitioners (Section 6.4.3). Finally, we used online surveys in order to understand and compare the perceived importance of these privacy attributes to experts and users (Section 6.4.4). The research methodology behind each of these three steps is described in more detail below.

6.3.1 Systematic review

The goal of the systematic review was to identify proposals from both academia and industry that can be used as sources of privacy attributes relevant for online services. We limited the scope of the review to documents which include either (a) an original visual representations of aspects related to privacy or data handling by online services or (b) a concrete list of high-level principles related to privacy or data handling for developing online services. While privacy is context-dependent, the goal of this study is to extract a general list of privacy attributes which are applicable to any kind of online service. Therefore, we are not interested in privacy attributes which are only relevant for a specific technology (e.g., mobile applications or IoT devices), domain (e.g., healthcare

or social networks), or specific target-group (e.g., children). Furthermore, we excluded anything published or last updated before 2001. This cut-off was chosen because 2001 marked a new climax in privacy concerns resulting from such developments as Web 2.0, and from the privacy vs. security debate ensuing the 9/11 terrorist attacks (H. J. Smith et al., 2011). We started by searching Scopus using the following queries, filtering out papers published before 2001.

- TITLE-ABS-KEY(privacy AND (label OR icons OR symbols)) resulting in total of 2063 papers;
- TITLE-ABS-KEY(“privacy by design“ AND (principles OR guidelines)) resulting in a total of 185 papers.

We then followed an iterative systematic review process (Siddaway et al., 2019), which is described in the remainder of this sub-section and visualized in Figure 6.1.

6.3.1.1 *Privacy visualizations*

We read the titles and abstracts of the 2063 papers retrieved from Scopus and identified 23 which might include an original visual representation of aspects related to privacy or data handling by online services. When scanning these papers, we learned of other terms used to describe privacy visualizations so we assembled these into an extended Scopus query, this time using phrases in order to reduce the amount of irrelevant results:

- TITLE-ABS-KEY (“privacy symbol” OR “privacy label” OR “privacy icon” OR “privacy graphic” OR “privacy visual” OR “privacy pictogram” OR “privacy indicator” OR “privacy indication” OR “privacy badge” OR “privacy emblem” OR “privacy image” OR “privacy motif” OR “privacy mark” OR “privacy token” OR “privacy stamp”) resulting in a total of 82 papers.

We read the titles and abstracts of these 82 results and identified ten more potentially relevant papers. We then read the full text of the 23+10 papers selected thus far and found 17 other relevant papers among their references, which were then also read. In total, we were able to find 41 papers containing privacy visualizations. We also learned about a 2016 Workshop on Privacy Indicators but found none of the papers published there satisfied our inclusion criteria. To make sure we did not miss anything, we performed several Google searches using all of the keywords we identified and found five more proposals for privacy visualizations coming from NGOs and industry.

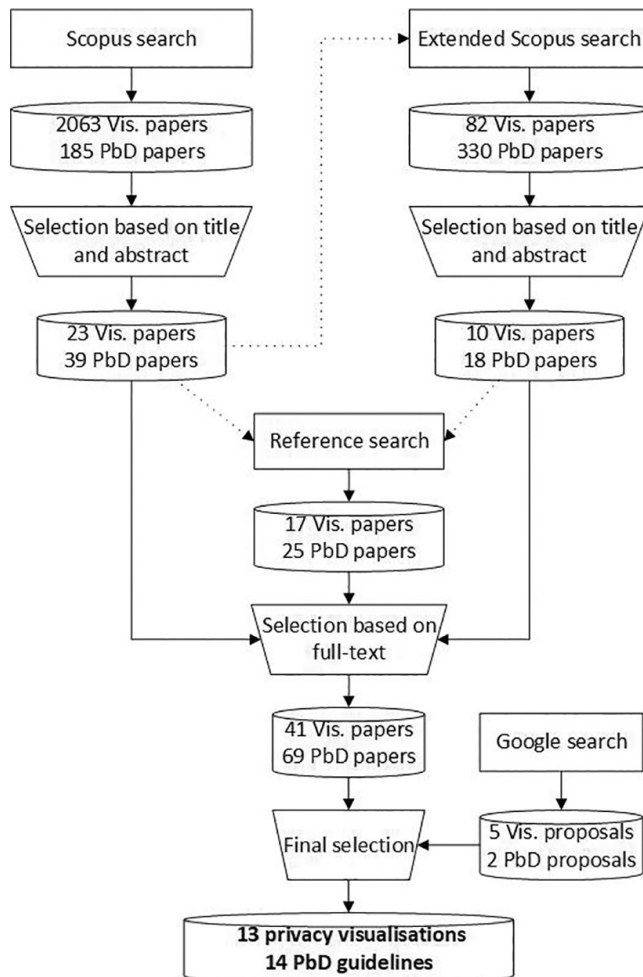


Figure 6.1 Systematic review process

Finally, we analyzed and discussed all of the 41+5 results in order to select suitable candidates for extracting privacy attributes applicable to online services in general. Below we briefly discuss the results which were excluded from our final sample.

Several proposed visualizations consist of ‘traffic-light’ style labels for websites (Cranor, 2009), hyperlinks (Hawkey & Inkpen, 2007), e-mails (Stojkovski & Lenzini, 2020), and search results (Zimmerman et al., 2019). Bal (2014) tested a privacy warning for app markets. These were excluded because they only include an overall rating and no individual attributes could be distinguished.

Quay-de la Vallee et al. (2016) propose displaying a rating next to each Android permission, Egelman et al. (2015) suggest icons indicating which apps have access to camera or microphone, and Levy and Gutwin (2005) insert warning icons next to input fields on websites. These were excluded because no individual privacy attributes could be distinguished.

Hansen (2009), Morel and Pardo (2020), Tsai et al. (2011) and Zorzo et al. (2016) evaluate and classify existing visualizations and were excluded because they do not offer any original visualizations.

Emami-Naeini et al. (2019) prototype a label to inform buyers of IoT devices, Gisch et al. (2007) proposes privacy badges that show the amount of data sent by a device, Johnson et al. (2020) discuss IoT security labeling and Y. Shen and Vervier (2019) adapts an existing privacy label to IoT devices. These papers were excluded because they are technology-specific.

Poneris et al. (2018) suggest pictograms for assistive technologies, De Lima Salgado et al. (2019) extend an existing label to smart toys, and K. L. Smith et al. (2017) report on co-designing icons for teens. Iannella and Finden (2009) evaluate privacy icons for photo sharing on social networks. These papers were excluded because the visualizations they proposed were target-group specific.

After removing duplicates, we ended up with a final sample of 13 privacy visualizations which we discuss in detail in Section 6.4.1 and which served as input for our coding process. Out of these 13, seven come from academia, five from industry, and one from government.

6.3.1.2 *Privacy by Design principles*

We read the abstracts and titles of the 185 papers retrieved from Scopus and selected 39 which appeared to include a concrete list of high-level principles related to privacy or data handling for developing online services. When scanning these papers, we learned of other related terms so we assembled these into an extended Scopus query:

- TITLE-ABS-KEY (“privacy by design” AND (principles OR guidelines OR conventions OR fundamentals OR rules OR strategies OR methods OR procedures OR protocols OR guide)) resulting in a total of 330 papers.

We read the titles and abstracts of these results and identified 18 more potentially relevant papers. We then read the full text of the 39+18 papers selected thus far and found 25 other relevant papers among their references, which were also read. In total, 69 papers containing high-level PbD principles

were found. We also ran a Google search using the original Scopus query and found two other proposals for PbD principles coming from industry.

We analyzed and discussed all of the 69+2 results in order to select which are suitable for extracting generally-applicable privacy attributes. Below we briefly discuss the results which were excluded from our final sample.

Many of the papers apply existing PbD principles to specific domains such as smart cities (I. Brown, 2014; Cavoukian, 2011a; Cavoukian & Chibba, 2016), smart factories (Preuveneers et al., 2016), healthcare (Bincoletto, 2019), design (Rostama et al., 2017), social networks (Vemou & Karyda, 2014), app development (Van der Sype & Maalej, 2014), blockchain (Barati & Rana, 2020), and ubiquitous systems (Davies & Langheinrich, 2013; Langheinrich, 2001).

Others propose technology-specific principles: Abdul-Ghani and Konstantas (2019), Perera et al. (2016), Perera et al. (2020) and the UK Government (Department for Digital Culture Media & Sport, 2018) ‘Secure by Design’ for IoT devices, Sedenberg et al. (2016) for robots, Pedraza et al. (2011, 2013) for facial recognition systems, Pinkas (2016) for eID systems, and Vanezi et al. (2019) for e-learning platforms. We excluded all of these from further analysis because we are looking for generally applicable privacy attributes.

Belli et al. (2017), Hansen et al. (2008), Makri and Lambrinouidakis (2015), Ringmann et al. (2018), Romanou (2018), and Schneider (2018) discuss and compare existing PbD principles. Tamburri (2020) and Tokas et al. (2020) formalize the PbD principles included in the GDPR. Several other papers (Ahmadian et al., 2019; Alshammari & Simpson, 2017; Baldassarre et al., 2019; Bier et al., 2014; Cavoukian, 2020; Colesky & Caiza, 2018; Colesky et al., 2018; Colesky et al., 2016; Finneran Denny et al., 2014; Drozd, 2016; Martín & Del Álamo, 2017; Mayfield, 2016; Suphakul & Senivongse, 2017) refine or implement PbD principles proposed by others. These were excluded because they do not provide any original principles.

After removing duplicates, we ended up with a final sample of 14 PbD guidelines which we discuss in detail in Section 6.4.2 and which served as input for our coding process. Out of these 14, two come from academia, five from industry, and seven from government.

6.3.2 Coding

To analyze the results of the systematic review, we followed an iterative coding process. First, the second author of this paper analyzed the privacy visualizations and PbD guidelines selected during the systematic review. The content was divided into passages and each passage was coded with one or

more terms related to the handling of personal data. This resulted in an initial list of 13 privacy attributes.

Second, we discussed the initial list of privacy attributes with two information security professionals from a large software solutions provider in the Netherlands in a 1-hour unstructured interview. Both security professionals deal with information privacy on a daily basis. The two experts were asked to check the list for completeness, to minimize overlap, to come up with unambiguous descriptions, and to ensure the attributes are both quantifiable and translatable into software requirements. As a result of the interviews, the initial list was refined by adding two new attributes. The experts also helped clarify the definitions of the attributes.

Third, to validate the refined list of 15 privacy attributes, three other coders coded independently 60% of the sample. After three rounds of discussions, refining the definitions of our codes, and re-coding the documents, Cohen's kappa reached .93, which indicates an almost perfect agreement between the coders and therefore validates our final list of attributes.

Fourth, the final list and corresponding description of attributes was used as a coding scheme for analyzing the full sample of 13 privacy visualizations and 14 PbD guidelines.

6.3.3 Online survey

To understand which attributes are most important, we designed an online survey to take the opinion of privacy experts and general users¹ (hereafter referred to as 'users') into account. A convenience sample of users was recruited via universities, online social networks, and two commercial subject pools. We recruited privacy experts via LinkedIn by first asking approximately 500 members with "privacy officer" in their profile description to connect. The ones that accepted the invitation were asked if they perceive themselves as suitable privacy experts for this study and—if so—were directed to the questionnaire.

The survey (approved by the ethical committee of the EEMCS faculty, University of Twente) collected demographic data about gender, education, occupation, nationality, and the type and frequency of online service usage. We asked the subjects how important on a scale from 0 (not at all important) to 10 (extremely important) they considered each of the 15 privacy attributes. Since the aim was to obtain an overall ranking, we did not select a specific scenario. To assess the sensitivity of the findings, we asked participants whether or not they

1 Within the realm of this dissertation, the definition of a general (or 'lay') user is a user who does not possess any specialized online privacy and/or cybersecurity expertise.

would rate these attributes differently for different types of services. Finally, in open questions, we asked if any of the descriptions were ambiguous and if they felt any attributes were missing.

By the 5th of December 2019, 646 adult participants (148 privacy experts and 498 users) had responded to the questionnaire. To clean the data, we removed all 86 incomplete responses from the sample. A further 75 responses were removed after being considered invalid due to: (1) questionable completion times (less than 2 minutes or more than 20 minutes), (2) pattern answering, (3) uncertainty—by their own admission—as to what the question was asking, or (4) no usage of online services. The number of valid responses was $N = 485$, of which 20.6% were privacy experts and 79.4% users. Of these, 49.7% were women and 48.9% men. European nationals made up 91.8% of the sample. All adult age groups were represented: 18-24 (35.1%), 25-34 (10.7%), 35-44 (13.2%), 45-55 (22.7%), and 55+ (15.9%). Many of the respondents were well educated, with either undergraduate degrees or post-graduate degrees (24.3% and 29.3% respectively). All respondents used online services at least once a day, and 66.2% did so several times a day. We also calculated the overall attribute importance, as the average score of the 15 privacy attributes. It was found to be reliable (Cronbach's $\alpha = 0.90$).

6.4 Results

6.4.1 Privacy visualizations

Privacy visualizations are visual representations designed to communicate aspects related to the handling of personal data to users of online services. In this section, we briefly describe the 13 privacy visualizations selected in chronological order and the privacy attributes they cover discussed.

6.4.1.1 Mehldau's data-privacy declarations

Mehldau was the first to propose an icon set to communicate the privacy aspects of an online service, in 2007. His list of 'data-privacy declarations' contained 30 icons grouped into four categories which could be used to represent how data are used, stored, shared or deleted. Because of the large number of icons, they are not shown here². The four categories were:

² The full list of icons is available under a CC-BY license from: <https://netzpolitik.org/wp-upload/data-privacy-icons-v01.pdf>.

- What data?, e.g., username, address, IP, contacts, cookies;
- How is my data handled?, e.g., deleted, saved, anonymized, encrypted, published;
- For what purpose?, e.g., statistics, advertising, shopping;
- For how long?, e.g., end of usage, timestamp, undetermined.

6.4.1.2 KnowPrivacy's Policy Coding Methodology

The KnowPrivacy research project³ aimed to create a set of coding tags in order to analyze privacy policies. In 2009 they proposed a set of tags for types of user data collected, general data practices and data sharing. Each tag consisting of an icon and a description, as shown in Figure 6.2. For each privacy policy, a tag could be in one of three states: YES, NO, or UNCLEAR (Gomez et al., 2009).














TYPE OF DATA COLLECTED	GENERAL DATA PRACTICES	DATA SHARING
 contact: name, mailing address, email, or phone number	 ad customization: user data may be used for the purpose of customizing advertising	 affiliates: affiliates and subsidiaries bound by the same privacy practices
 computer: IP address, browser type, or operating system	 third party tracking: site allows third parties to place advertisements that may track user behavior	 contractors: third party contractors bound by the same privacy practices
 interactive: browsing behavior or search history	 public display: service allows users to contribute information which may be displayed publicly	 third parties: third parties not subject to same data practices
 financial: account status or activity, credit information, or purchase history	 user control: users allowed to access and correct personal data collected	
 content: contents of personal communications, stored documents or media	 data retention: explicitly stated duration of retention for personal data collected	

Figure 6.2 KnowPrivacy's Policy Coding Methodology

3 <http://knowprivacy.org>

6.4.1.3 CyLab's privacy nutrition label

Developed by Carnegie Mellon's CUPS (CyLab Usable Privacy and Security) laboratory in 2009, the privacy nutrition label takes a tabular approach to represent how personal data are handled by an e-service provider (Kelley et al., 2009; see Figure 6.3). Each row corresponds to a data item (e.g., location, health information, etc.) and each column corresponds to a way in which each item is used (e.g., marketing, profiling, sharing with other companies, etc.). Each cell in the resulting matrix gives a visual indication with regard to each data item - usage pair:

- An exclamation mark on a dark or red background signifies that the item is used for that purpose;
- the text OUT on a dark gray or light red background signifies that the item is used for that purpose unless the user opts-out;
- the text IN on a light gray or dark blue background signifies that the items is not used for that purpose unless the user opts-in;
- a dash on a light background signifies that the data item is neither collected nor used for that purpose.

The rows and columns are fixed so that two policies can be compared side-by-side. There are a total of ten data items and seven ways in which these can be used. The possible usages are: (1) *provide service and maintain site*, (2) *research and development*, (3) *marketing*, (4) *telemarketing*, (5) *profiling*, (6) *sharing with other companies*, and (7) *sharing on public forums*.

Bell Group

information we collect	ways we use your information				information sharing	
	to provide service and maintain site	marketing	telemarketing	profiling	other companies	public forums
contact information		opt in			opt out	
cookies						
demographic information		opt in			opt out	
financial information						
health information						
preferences						
purchasing information		opt in			opt out	
social security number & gov't ID						
your activity on this site		opt in			opt out	
your location						

Figure 6.3 CyLab's privacy nutrition label example

6.4.1.4 Mozilla's privacy icons

In 2010, Aza Raskin from Mozilla proposed a set of icons which could be attached to existing privacy policies in order to provide a visual summary of the most important privacy issues: *retention period*, *third-party use*, *ad networks*, and *law enforcement* (Moskowitz & Raskin, 2011). The icon designs to represent these attributes have been the subject of multiple iterations, the latest are show in Figure 6.4. The project has since been abandoned but the icons are still present on Mozilla's Wiki⁴.

4 https://wiki.mozilla.org/Privacy_Icons

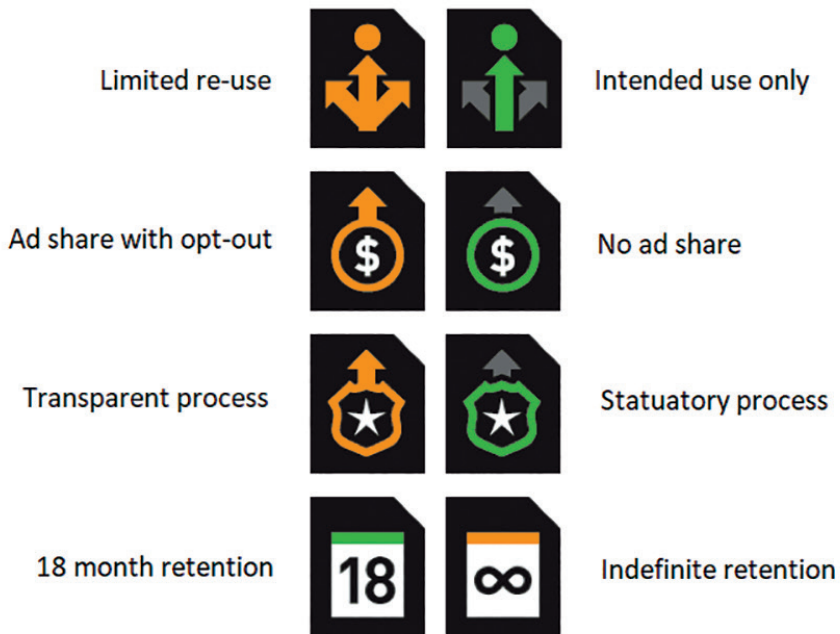


Figure 6.4 Mozilla's privacy icons v2

6.4.1.5 The PrimeLife project

Also in 2010, the EU-funded PrimeLife project⁵ published several sets of icons: a general set and other sets for specific domains such as social media (Holtz et al., 2011a). The icons were designed to be aligned with European privacy laws. The initial proposal contained 30 icons representing three types of privacy concepts: *data types* (i.e. personal, sensitive, payment, or medical data), *data purpose* (i.e. legal obligation, shipping, tracking, or profiling), and *data processing* (storage, deletion, pseudonymization, anonymization, disclosure, and collection). For social networks, PrimeLife added icons for groups of recipients (friends, friends of friends, selected individuals, and public). They performed user studies to compare different designs and found that icons should be as simple as possible and culturally-neutral, and their number held to a minimum. Figure 6.5 shows the icons rated highest during their evaluation.

5 <http://primelife.ercim.eu>



Figure 6.5 Highest rated icons from the PrimeLife project

6.4.1.6 TrustArc’s Privacy Short Notice

In 2011, TrustArc—the developers of the TRUSTe privacy certification standard—proposed an icon-based ‘privacy short notice’, aimed at providing a simplified summary of privacy policies (see Figure 6.6). After analyzing previous approaches, they concluded that such a short notice should focus on the data practices and uses that are invisible to users, namely *secondary use* (none, customization, or profiling), *sharing* (none, affiliates, or unrelated), *third-party tracking*, and *data retention* (none, limited, or indefinite; Pinnick, 2011).



Figure 6.6 Icons of the TrustArc short notice

6.4.1.7 Privacy wheel

Based on survey results that revealed users’ preference for general and less legally detailed information about data handling practices, Van den Berg and Van der Hof (2012) developed the privacy wheel. Taking OECD’s privacy principles as a starting point for their visualization, the wheel (see Figure 6.7) covers eight core concepts of privacy related information: (1) *collection*, (2) *data quality*, (3) *purpose*, (4) *limited use*, (5) *security*, (6) *consent*, (7) *third parties*, and (8) *accountability*. The spokes of the wheel are clickable, providing two layers of increasingly detailed information. Furthermore, some spokes provide an interactive mechanism for updating opt-in/opt-out preferences.



Figure 6.7 The Privacy Wheel

6.4.1.8 GDPR’s draft privacy icons

Article 12 of the European GDPR mandates that “the information may be provided in combination with standardized icons in order to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing” (The European Parliament and the Council of European Union, 2016). In addition, it specifies that the icons should be machine readable. The final version of the GDPR does not prescribe specific icons or attributes which need to be represented, but empowers the European Commission to determine these at a later time. However, an earlier draft of the GDPR did explicitly describe six icons shown in Figure 6.8. For each icon, a given application may score a checkmark or an X.







ICON	ESSENTIAL INFORMATION	FULFILLED
	No personal data are collected beyond the minimum necessary for each specific purpose of the processing	
	No personal data are retained beyond the minimum necessary for each specific purpose of the processing	
	No personal data are processed for purposes other than the purposes for which they were collected	
	No personal data are disseminated to commercial third parties	
	No personal data are sold or rented out	
	No personal data are retained in unencrypted form	

Figure 6.8 Privacy icons from GDPR draft

6.4.1.9 DCIs

Developed in 2017 by researchers from the University of Oxford and Cambridge, the Data Controller Indicators provide information on the kind of data that are sent by an app to various parties while considering the background information of those parties and the purposes behind data usage (Van Kleek et al., 2017). The personalized version of DCIs (PDCIs) even goes a step further by integrating third-party libraries (see Figure 6.9). Testing different versions of the visualization with users revealed a preference for the personalized DCIs which provides a differential risk assessment of data controllers (if those providers already accessed data via other apps or if data are newly disclosed to an organization).



Figure 6.9 Example DCIs label

6.4.1.10 Fox et al.'s GDPR compliant label

In 2018, Fox et al. (2018) started developing a privacy label that is compliant to the requirements mandated in the GDPR. Their label is based on the CyLab's privacy nutrition label. In an iterative process, the authors developed an icon- and a text-based label and tested them in the context of an e-commerce website, revealing users' preference for the icons. Consequently, an icon-based label was further developed, covering twelve privacy attributes as shown in Figure 6.10: (1) *information about data controller*, (2) *data processing purposes*, (3) *recipients of personal data*, (4) *transfer to third countries*, (5) *retention*, (6) *rights of the data subject*, (7) *consent*, (8) *right to complain*, (9) *disclosure*, (10) *automated decision-making*, (11) *details of data protection officer*, and (12) *further data processing*. Furthermore, a toggle on/off function had been added to the label. As this project is still in progress, the authors aim for testing the label with and without explicit consent function and under consideration of the Mobile Privacy-Security Knowledge Gap model (Crossler & Bélanger, 2017).

6.4.1.11 CLEVER°FRANKE's privacy label

In 2019, SensorLabs⁶, a Dutch non-profit initiative by UX design firm CLEVER°FRANKE published a highly simplified privacy label (Franke et al., 2019). The label is designed for online services as well as physical devices. They envisioned the label being attached to 'things' such as vending machines, card scanners and even storefronts. To come up with the label, they reviewed the literature on conceptualizing and extracted three essential aspects of privacy, namely (1) *collection*, (2) *purpose* and (3) *control*. Each of these aspects is measured using five yes/no questions based on the Rathenau Institute's overview of ethical and societal issues related to digitization (Kool et al., 2017). Each 'yes' answer achieves one point, up to a maximum of 15 points. The final score determines what label the entity receives. Each label consists of two elements: an


6 <https://www.sensorlab.nl/research-program/>

A-to-F category which also determines the color (A is green, F is red, everything in-between is shades of orange), and a visual representation of the score on each of the three aspects. Figure 6.11 shows some example labels. The circle around the letter is divided into three parts—corresponding to collection, purpose, and control. Each part consists of five layers—corresponding to the five questions for each aspect.


To Your Door Store Privacy Policy

512 Pearse Street, Dublin 2, Ireland
info@toyourdoorstore.eu +353 1 202 786


Information We Collect




Basic Information
i.e. email address, name and physical address



Financial Information
purchasing data and credit card details




Location Information
Your location and IP Address




Usage Information
We use cookies to track the products and pages you view


How We Use The Information We Collect




Marketing
To market new products and special offers



Transactions
We save your data to speed up future transactions



Encryption ⓘ
Information is encrypted and stored securely



Data Storage ⓘ
We do not transfer any of your information outside the EU

Your Rights

You can:

- Update your consent settings at any time by toggling on the options above
- Request a copy of your data in your account settings
- Update or delete your information in your account settings

Information on your additional rights including how to complain to the supervisory authority can be found in our [full privacy policy](#).

For more information

Contact our Data Protection Officer Marie Byrne by email at maria@toyourdoorstore.com or by phone +353 1 202 786

Figure 6.10 Example of Fox et al.'s GDPR compliant label



Figure 6.11 Five example labels from CLEVER°FRANKE

6.4.1.12 DaPIS

The Data Protection Icon Set, developed by Rossi and Palmirani in 2019, is characterized by a multilayered structure and based on PrOnto, a computational ontology of the GDPR. The machine-readable layers provide interpretable information from legal documents whereas the human-centered layer adds visual accessible icon design. DaPIS covers six main classes (see Figure 6.12): (1) *data* e.g., personal, (2) *agents' roles* e.g., the data subject or controller, (3) *processing operations*, e.g. anonymization or profiling, (4) *the data subject's rights*, e.g., access or erasure, (5) *processing purposes*, e.g., research or marketing, and (6) *legal bases for processing*, e.g., consent or legitimate interest. The authors emphasize that DaPIS is not designed to be a standardized European icon set but provides a foundation for the implementation of GDPR's icons and is still under development.

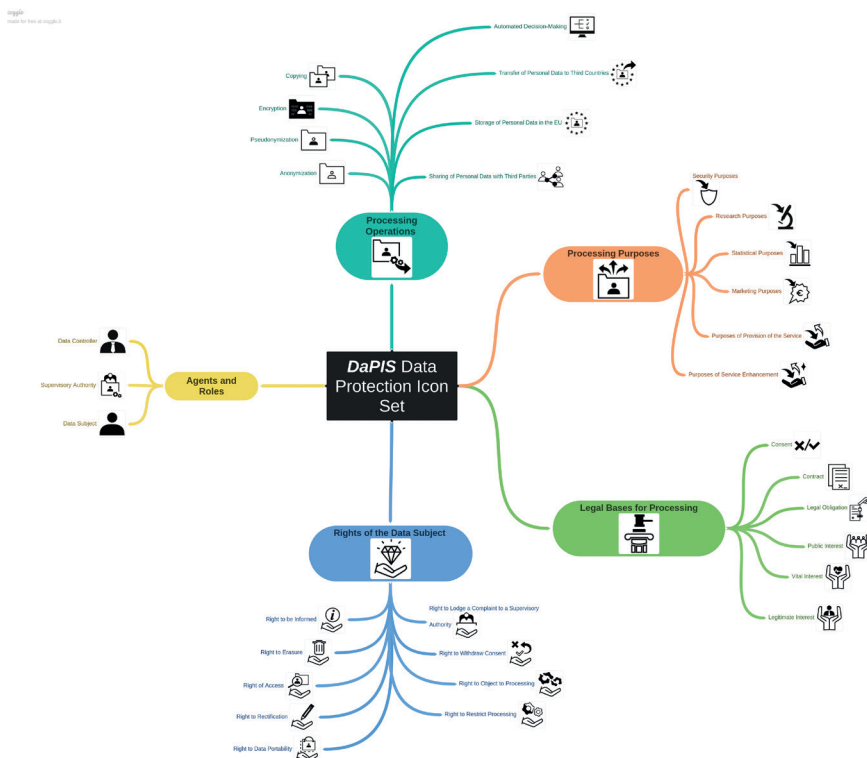


Figure 6.12 The Data Protection Icon Set (DaPIS)

6.4.1.13 Privacy Label

In 2020, a Dutch consortium of privacy related companies and non-profit organizations, launched Privacy Label⁷ that summarizes the privacy statement of an online service. In combination with graphical icons, the tabular label provides information on seven core themes: (1) *collection*, (2) *purpose*, (3) *sharing*, (4) *location*, (5) *duration*, (6) *legal basis*, and (7) *take action*. See Figure 6.13 for an example. The title of each theme is clickable, providing short explanations in relation with the GDPR regulations. For more information, a ‘learn more’ option is provided; directing the user to the Privacy Label website. Furthermore, each core theme contains several sub-themes (referred to as ‘ingredients’) that are, in turn, clickable and provide further information tailored to the data practices of the online service (e.g., the reason behind data aggregation).

Privacy Label of Example webshop This Privacy Label is the general label of our example webshop. This basically is a summary of its privacy statement.	
 Data collection ⓘ We receive from you: personal data ⓘ We receive from others: aggregated data, personal data & sensitive personal data ⓘ We observe: personal data & sensitive personal data ⓘ Derive: personal data & sensitive personal data ⓘ	 Location ⓘ Most data is processed outside the EU ⓘ  Duration ⓘ Most data: one month or less Some data: one year or less A little data: ten years or less
 Purpose ⓘ Providing goods and services ⓘ Automated decision making ⓘ Marketing, sales and customer relationship ⓘ Financial administration ⓘ	 Legal grounds ⓘ Contract ⓘ Legitimate interest ⓘ Legal obligation ⓘ Consent ⓘ
 Data sharing ⓘ Service providers ⓘ Partners ⓘ Processors ⓘ Advertisers ⓘ Government ⓘ	 Take action ⓘ Read our privacy statement Manage your data Contact our privacy representative Email: info@privacylabel.org Phone: 0612345678
Last updated 2020-04-24	Privacy Label Version 2020-04 CODE

Figure 6.13 Example of a label generated by Privacy Label^{6.4.2 Privacy by Design guidelines}

7 privacylabel.org; created and supported by ECP, PineappleJazz, Privacy Company, SURF, and SIDN funds

6.4.2 Privacy by Design guidelines

Privacy by Design (PbD) is an umbrella term for software development approaches that take privacy considerations into account from the early stages of design. In this section, we briefly describe each of the 14 PbD guideline selected in chronological order and summarize the principles it proposes.

6.4.2.1 The Australian Privacy Principles

The Australian Privacy Principles (APPs) were first added to the Australian Privacy Act in 2001. The APPs apply to the private sector and most government entities in Australia. They are technology neutral, and can be tailored to the needs of individual organizations. In 2014, the original list of ten principles was extended to 13 (Office of the Australian Information Commissioner, 2014):

- *Open and transparent management of personal information:* Manage personal data in an open and transparent way, including a clear and up-to-date privacy policy.
- *Anonymity and pseudonymity:* Provide individuals the option of not identifying themselves.
- *Collection of solicited personal information:* Conditions for collecting personal or sensitive data when needed and allowed.
- *Dealing with unsolicited personal information:* Avoid gathering unsolicited personal data.
- *Notification of the collection of personal information:* Provide information about data collection.
- *Use or disclosure of personal information:* Conditions for usage or disclosure of personal data.
- *Direct marketing:* Restrict use, disclosure of personal data for direct marketing purposes.
- *Cross-border disclosure of personal information:* Conditions for personal data protection before disclosure overseas.
- *Adoption, use or disclosure of government related identifiers:* Conditions for government related identifier adoption, or the disclosure of it.
- *Quality of personal information:* Ensure personal data collected, used, or disclosed is accurate, up-to-date, and complete.

- *Security of personal information*: Protect personal data and remove them when needed.
- *Access to personal information*: Conditions for providing access to personal data.
- *Correction of personal information*: Obligations for amendment of personal data.

6.4.2.2 CSA's Model Code for the Protection of Personal Information

Firstly published in 1996, the Canadian Standards Association (CSA Group, 2014) has reaffirmed the Model Code for the Protection of Personal Information in 2001. The standard is focused around privacy rights and individual control over the use and exchange of personal information. Eventually, the ten principles developed by the CSA have been incorporated into Canadian law. The following principles form the basis of the Model Code for the Protection of Personal Information:

- *Accountability*: Responsibility for personal data and compliance with the principles.
- *Identifying purposes*: Identification of purposes before or at the time of data collection.
- *Consent*: Consent is required for the collection, use, or disclosure of personal data.
- *Limiting collection*: Data collection is limited to specified purposes.
- *Limiting use, disclosure and retention*: Disclosure and retention limited to purposes.
- *Accuracy*: Accuracy, completeness and up-to-dateness of personal data.
- *Safeguards*: Data protection in proportion to the sensitivity of the information.
- *Openness*: Readily available privacy policies and data management information.
- *Individual access*: Upon request, access and amendment of personal data.
- *Challenging compliance*: Possibility to challenge compliance with the principles.

6.4.2.3 APEC's Privacy Framework

Published in 2005, Asia-Pacific Economic Cooperation (APEC) developed a principle-based privacy framework. Inspired by OECD guidelines, this framework aims at developing information privacy protections and to warrant the free information flow in the Asia Pacific region. The privacy framework includes the following privacy principles:

- *Preventing Harm*: Prevention of misuse of personal information.
- *Notice*: Provision of clear and easily accessible privacy policies.
- *Collection Limitation*: Limitation of information collection to purpose.
- *Uses of Personal Information*: Usage of personal data limited to purposes.
- *Choice*: Possibility to exercise choice regarding collection, use, and disclosure of data.
- *Integrity of Personal Information*: Accuracy, completeness, and up-to-dateness of data.
- *Security Safeguards*: Protection of data against risks, e.g., loss or unauthorized access.
- *Access and Correction*: Provision of access to personal data and the ability to correct them.
- *Accountability*: Responsibility for compliance with these principles.

6.4.2.4 The Global Privacy Standard

The Global Privacy Standard (GPS), was published in 2006, at the 28th International Data Protection and Privacy Commissioners Conference. Its purpose was to reinforce the mandate of data protection authorities by drafting 'fundamental and universal privacy concepts' (Cavoukian, 2006), namely:

- *Consent*: Consent for collection, use or disclosure of personal information, and ability to withdraw consent.
- *Accountability*: Communicate all privacy policies and procedures and seek equivalent privacy protection from third parties.
- *Purposes*: Specify and communicate the purpose for collecting, using, retaining and disclosing personal information.

- *Collection Limitation and Data Minimization*: Collection is fair, lawful and limited to specified purposes; data minimization and anonymization or pseudonymization should be applied.
- *Use, Retention, and Disclosure Limitation*: Limit use, retention, and disclosure of personal information to specified purposes, except when required by law.
- *Accuracy*: Accurate, complete, up-to-date personal information as per the specified purposes.
- *Security*: Ensure security of personal information throughout its lifecycle as per recognized international standards.
- *Openness*: Make information about policies and practices related to personal information readily available.
- *Access*: Provide access to personal information, its uses, and allow to challenge its completeness or have it amended.
- *Compliance*: Monitor, evaluate, and verify compliance with privacy policies and procedures.

6.4.2.5 ISTPA's Privacy Framework

In 2007, triggered by considerable changes in information privacy since 2002, as well as huge variations in the language and content of existing privacy frameworks, the International Security, Trust and Privacy Alliance (ISTPA) performed a structured review of existing privacy regulations and standards and extracted a set of key principles. They supplemented the list with three additional principles, resulting in a working set of 11 privacy principles (Sabo, 2007):

- *Notice*: Provision of an overarching privacy policy.
- *Consent*: Opt-in/opt-out, or implied affirmative process.
- *Collection Limitation*: Minimal data collection and related to purposes.
- *Use Limitation*: Usage and retention of personal data for specified purposes only.
- *Disclosure*: Release, transfer, access or re-use of data with consent of the data subject only.
- *Access and Correction*: Ability to access and amend personal data.
- *Security/Safeguards*: Confidentiality, availability and integrity of personal data.

- *Data Quality*: Adequacy, up-to-dateness, minimization or elimination of personal data in relation to purposes.
- *Enforcement*: Assurance of compliance with privacy policy and ability to challenge this.
- *Openness*: Availability of privacy policy.
- *Anonymity*: Prevention of identification.
- *Data Flow*: Communication of data across geo-political jurisdictions.
- *Sensitivity*: Specification of data that need special security controls.

6.4.2.6 The Generally Accepted Privacy Principles

In 2009, the American Institute of Certified Public Accountants and the Canadian Institute of Chartered Accountants (AICPA & CICA, 2009) published GAPP (Generally Accepted Privacy Principles). It was intended as a global privacy framework aimed at helping accountants develop their own privacy program. GAPP is supported by 70 objectives grouped under ten core principles:

- *Management*: Communicate, and assign accountability for privacy policies and procedures.
- *Notice*: Notice about privacy policies and procedures, identify purposes for personal information collection, usage, retention, and disclosure.
- *Choice and consent*: Describe the choices and obtain implicit or explicit consent for the collection, use, and disclosure of personal information.
- *Collection*: Collect personal information only for the purposes identified in the notice.
- *Use, retention, and disposal*: Limit use and retention of personal information to identified and consented purposes or as required by law and thereafter disposal of such information.
- *Access*: Provide individuals with access to their personal information for review and update.
- *Disclosure to third parties*: Disclose personal information to third parties only for the purposes identified in the notice and with the implicit or explicit consent of the individual.
- *Security for privacy*: Protect personal information against unauthorized access (both physical and logical).

- *Quality*: Maintain accurate, complete, and relevant personal information for the purposes identified in the notice.
- *Monitoring and enforcement*: Monitor compliance with privacy policies and procedures and have procedures to address privacy-related complaints and disputes.

6.4.2.7 Cavoukian's 7 Foundational Principles

Also in 2009, Ann Cavoukian (2009, 2010, 2011b), the Privacy Commissioner of Ontario combined Langheinrich's Principles of Privacy-Aware Ubiquitous Systems (Langheinrich, 2001) with those of the GPS (Cavoukian, 2006) into a set of high-level design principles for privacy-aware software that were later adopted by Deloitte (Casey, 2016):

- *Proactive not Reactive; Preventive not Remedial*: Anticipate and prevent privacy-invasive events instead of resolving them after they occur.
- *Privacy as Default*: Ensure personal data protection automatically, without requiring action from individuals.
- *Privacy Embedded into Design*: Embed PbD into the design by making it a core functionality and not an add-on.
- *Full functionality - Positive-Sum, not Zero-Sum*: Accommodate all legitimate interests and avoid unnecessary trade-offs and false dichotomies such as privacy vs. security.
- *End-to-End Life-cycle Protection*: Secure data from start to finish and ensure they are securely destroyed at the end of the process.
- *Visibility and Transparency*: Assure stakeholders that data are handled in accordance with stated promises and objectives and ensure visibility and transparency.
- *Respect for User Privacy*: Protect the interests of the individuals by offering strong privacy defaults, appropriate notice, and by empowering user-friendly options.

6.4.2.8 ISO29100 Privacy Framework

In 2011, the ISO/IEC Information Technology Task Force (ITTF) published its own privacy framework specifying a common privacy terminology while defining actors and roles involved in the processing of Personally Identifiable Information

(PII; Technical Committee ISO/IEC JTC 1/SC 27, 2011). Revised in 2017, the standard defines its own set of privacy safeguarding considerations, namely:

- *Consent and choice*: Inform PII principals about PII processing, their rights, available choices, and implications; obtain consent and allow it to be withdrawn easily and free of charge.
- *Purpose legitimacy and specification*: Ensure purpose(s) comply with law; communicate purpose(s) to PII principals before the time the information is collected or used for a new purpose.
- *Collection limitation*: Limit the collection of PII to the bounds of applicable law and strictly necessary for the specified purpose(s).
- *Data minimization*: Minimize the amount of PII processed and the number of third-parties involved, strive for anonymity or pseudonymity and delete PII when retention is no longer necessary.
- *Use, retention and disclosure limitation*: Limit use, retention and sharing of PII to the purposes specified.
- *Accuracy and quality*: Ensure that PII processed is reliable, accurate, complete, up-to-date, and periodically check and verify the validity and correctness before making any changes.
- *Openness, transparency, and notice*: Provide clear and accessible information about policies and procedures concerning PII process, review, and correction and notice about any major changes.
- *Individual participation and access*: Provide PII principals with the ability to access and review PII, to challenge accuracy, have it amended, corrected or removed without cost or delay.
- *Accountability*: Document and communicate privacy policies and procedures; define complaint procedures, inform about privacy breaches, including sanctions and compensation.
- *Information security*: Protect PII with controls at the operational, functional and strategic levels to ensure integrity, confidentiality, and the availability of PII throughout its life-cycle.
- *Privacy compliance*: Have appropriate internal controls and independent supervision mechanisms, periodically conduct audits perform privacy risk assessments.

6.4.2.9 OECD's privacy principles

Based on their 1980 Fair Information Practices aimed at the trans-border flow of information, the Organisation for Economic Cooperation and Development (OECD) published a revised set of privacy principles in 2013 which integrated the recent work on privacy law enforcement cooperation, resulting in the following principles (OECD, 2013):

- *Collection Limitation Principle*: Limited, fair, lawful data collection, obtain informed consent.
- *Data Quality Principle*: Keep personal data relevant, accurate, complete, and up-to-date.
- *Purpose Specification Principle*: Specify intended use before collection.
- *Use Limitation Principle*: Do not use personal data for purposes other than those specified.
- *Security Safeguards Principle*: Protect personal data using reasonable security safeguards.
- *Openness*: Be transparent about the handling of personal data and provide contact information.
- *Individual Participation*: Provide easy access to personal data and the ability to remove them.
- *Accountability Principle*: Be accountable for complying with the principles stated above.

6.4.2.10 Hoepman's Privacy Design Strategies

First published in 2014, Hoepman's 'Little Blue Book' outlines PbD strategies attempting to make PbD more concrete and applicable in practice (Hoepman, 2014). The book translates legal norms and best-practices surrounding personal data into the following design requirements:

- *Minimise*: Keep the amount of personal information processed to a minimum.
- *Hide*: Hide any personal information that is processed from plain view.
- *Separate*: Process personal information in a distributed fashion whenever possible.
- *Aggregate*: Process personal information at the highest aggregation level and with the least detail.

- *Inform*: Inform data subjects adequately whenever personal information is processed.
- *Control*: Provide data subjects with agency over the processing of their personal information.
- *Enforce*: Have a privacy policy compatible with legal requirements in place and enforce it.
- *Demonstrate*: Demonstrate compliance with privacy policy and legal requirements.

6.4.2.11 OASIS Privacy Management Reference Model

The Privacy Management Reference Model and Methodology (PMRM) was developed and published in 2016 by the Organization for the Advancement of Structured Information Standards (OASIS), a non-profit organization committed to privacy and personal data protection. Derived from international legislation and regulations, the PMRM provides a set of 14 privacy principles (Drgon et al., 2016):

- *Accountability*: Compliance with privacy policies.
- *Notice*: Open and transparent privacy policies.
- *Consent and Choice*: Opt-in/opt-out, or implied affirmative process.
- *Collection Limitation and Information Minimization*: Data collection, processing and retention limited to purpose fulfillment.
- *Use Limitation*: Usage limited to specified and accepted purposes.
- *Disclosure*: Transfer, access, or re-use of personal data with consent permission.
- *Access, Correction and Deletion*: Right to discover, correct or delete personal data, right to be forgotten.
- *Security/Safeguards*: Confidentiality, availability and integrity of personal data.
- *Information Quality*: Accuracy, correctness and up-to-dateness of personal data.
- *Enforcement*: Compliance with privacy policies.
- *Openness*: Access to information about data handling practices.
- *Anonymity*: Prevention of identification.

6.4.2.12 The Privacy Company's PbD Framework

In 2018, the Privacy Company published a data protection by design framework aimed at developers (The Privacy Company B.V., 2019). It translates the requirements of the European GDPR into the following guidelines:

- *Anonymization*: Anonymize and aggregate.
- *Data minimization*: Gather only necessary data and delete unnecessary data immediately.
- *Pseudonymization*: Remove directly identifying elements, hashing, polymorphic pseudo-ID.
- *Encryption*: Use public-key encryption, disk encryption, etc.
- *Access control*: Use digital data vault, logical access controls, authentication and authorization.
- *Data protection by default*: Provide privacy-friendly settings by default, transparent user interface, and permission management.
- *Deletion/Retention terms*: Automate deletion, data 'flagging' after end of retention term, sticky policies, data fading.
- *Facilitate rights of data subjects*: Privacy dashboard, communication/support.

6.4.2.13 GDPR Art. 5

Launched in 2018, the European General Data Protection Regulation (GDPR) regulates data privacy laws across Europe and replaced the Data Protection Directive 95/46/EC. All organizations that target or collect data from people within EU must comply with the GDPR. Article 5 of the GDPR covers the following seven data protection principles relating to the processing of personal data (The European Parliament and the Council of European Union, 2016):

- *Lawfulness, fairness and transparency*: Lawful, fair and transparent processing.
- *Purpose limitation*: Specification of legitimate purposes for data processing.
- *Data minimization*: Collection and processing of data restricted to what is absolutely necessary.
- *Accuracy*: Data kept accurate and up-to-date.
- *Storage limitation*: Storage only as long as necessary for purpose fulfillment.

- *Integrity and confidentiality*: Appropriate security, integrity, and confidentiality.
- *Accountability*: Responsibility for compliance with these principles.

6.4.2.14 *The Personal Information Protection and Electronic Documents Act*

Under the authority of the Office of the Privacy Commissioner of Canada, the Personal Information Protection and Electronic Documents Act (PIPEDA) was revised in 2019. PIPEDA provides ten fair information principles that serve as the groundwork for the collection, use, disclosure of, and access to personal data handled by the private sector (Office of the Privacy Commissioner of Canada (OAIC), 2000):

- *Accountability*: Responsibility for personal data and compliance with principles.
- *Identifying Purposes*: Specification of purposes before or at the time of collection.
- *Consent*: Collection, usage, or disclosure of personal data with consent.
- *Limiting Collection*: Limitation of collection according to purposes.
- *Limiting Use, Disclosure, and Retention*: Usage or disclosure only for specified purposes and retention limited to purpose fulfillment.
- *Accuracy*: Accuracy, completeness, and up-to-dateness of data.
- *Safeguards*: Appropriate security measures and in accordance with data sensitivity.
- *Openness*: Publicly and readily available privacy policy.
- *Individual access*: Provision to access data and ability to challenge accuracy and completeness.
- *Challenging Compliance*: Ability to challenge compliance with principles.

6.4.3 **Unified list of privacy attributes**

By means of an open-coding procedure, we distilled an initial list of 13 privacy attributes from the privacy visualizations and PbD guidelines we reviewed: accountability, anonymization, collection, control, correctness, disclosure, functionality, purpose, retention, sale, security, sharing, and transparency. After discussing this list with two practitioners (see Section 6.3.2), we added pseudonymization and the right to be forgotten before adding simple definitions to each of the privacy attributes. Finally, we iteratively refined the definitions

during rounds of coding, arriving at the following unified list of privacy attributes, ordered alphabetically:

Accountability = Can the service provider be held accountable for violations? *e.g., legally binding privacy policy, legal precedents, regulation, etc.*

Anonymization = Are all identifiable markers completely removed so that data can never be traced back to a single person? ^{*8}

- High level data aggregation is part of anonymization.

Collection = What data are collected? *e.g., IP address, phone number, credit card information, etc.*

- A major distinction can be made between personal data (information that relates to an identified or identifiable living individual) and anonymous data. Further distinction can be made between various types of personal data.
- Data minimization is part of collection: Collect as little data as possible; only data that are needed for provision of the service.

Control = Is the data subject able to choose or decide which data to share and for which purpose, and how difficult is it to do so?

- The core element of control is a self-determined decision on what to share and/or for which purpose and is the user able to actively influence how the service provider handles their personal data?
- Control includes obtaining informed consent as well as the ability to request a copy of the data—and is directly related to the user-friendliness of privacy settings.

Correctness = Are there mechanisms for preventing and fixing incorrect data? *e.g., data request forms, ability to edit collected data, etc.*

- Correctness has to do with the ability of the service provider and/or is the user able to fix incorrect data after the data were collected?

8 * Rephrased in the online survey for simplicity

- Correctness goes a step further than control: If data are already disclosed, is the user able to correct data about themselves which are not (or no longer) valid?

Disclosure = What is the attitude of the service provider toward requests from law enforcement? *e.g., disclosure upon request, disclosure only with a warrant, disclosure only after court order, etc.*

- Disclosure is about how the service provider reacts to requests from government institutions and concerns the jurisdiction of where data are stored or processed, *e.g., data leaving the EU.*

Functionality = Is the user forced to choose between functionality and privacy? *e.g., application does not run without accepting all permissions, only real names allowed, credit card details required for free trial, etc.*

- Functionality is about whether the service provider artificially restricts the service or parts of the service unless personal data are provided.

Purpose = What are the collected data used for? *e.g., provision of the service, advertising, profiling, etc.*

- Purpose includes the legal basis for processing (*e.g., data collected because of legal requirements or for vital/public interest*).

Pseudonymization = Are personally identifiable markers replaced by artificial identifiers, or pseudonyms, such that data can only be traced back to individual users with the help of additional information?* *e.g., names replaced by numbers, house number removed from address, birthday replaced by birth year, etc.*

Retention = How long are the collected data stored?

Right to be forgotten = Can data subjects request that all personal data are removed?

- Implementation can vary between hiding personal data and completely removing personal data.

Sale = Are any of the data sold to third parties?

- Sale has to do with obtaining commercial gains by sharing user data with other organizations.

Security = What technical measures are taken to ensure that data are protected from unauthorized or malicious access?

Sharing = Do any of the collected data leave the ownership of the service provider? *e.g., other companies, advertisers, research institutions, etc.*

- Sharing is sometimes referred to as disclosure—and includes both voluntary and unintentional disclosure of data.
- Sharing refers to data shared without monetary compensation.

Transparency = Is the user able to obtain information with regard to how their personal data are handled? *e.g., open-source code, availability of privacy policy, regular audits, etc.*

- Transparency includes clarification before giving informed consent or, in other words, proactive distribution of information to the user.
- Transparency is about whether the service provider can adequately demonstrate the implementation of all the other privacy attributes on this list to data subjects and regulators.

Table 6.1 shows which attributes were covered by privacy visualization or PbD guidelines. Notably, most privacy visualizations and PbD guidelines cover issues regarding collection and purpose. However, data sharing is only covered by half of the PbD guidelines. Furthermore, while all PbD guidelines make statements about security and transparency requirements, only half of the privacy visualizations we reviewed communicate these aspects to users. Accountability and correctness are also mentioned frequently in PbD principles but were rarely covered by privacy visualizations. Functionality was only found in Cavoukian’s PbD guidelines (Cavoukian, 2009, 2010, 2011b) and CLEVER°FRANKE’s privacy label (Franke et al., 2019), and sale of data is only covered by two privacy visualizations and zero Pbd guidelines. The similarities and differences are discussed in detail in Section 6.5.

6.4.4 Perceived importance of privacy attributes

By means of the online survey described in Section 6.3.3, 385 users and 100 privacy experts ranked the importance of the privacy attributes as described in Section 6.4.3. Figure 6.14 shows the mean importance of each attribute for the users and the privacy experts as well as the 95% confidence intervals. First we summarize the most important differences and similarities between users and privacy experts.

- Both users and privacy experts in our study agree that collection, sharing, and sale are the most important privacy attributes.
- Privacy experts assign up to about 10% more importance than users to most attributes. On the other hand, the same experts assign anonymization and the right to be forgotten with up to 10% less importance than users.
- The mean scores of users and privacy experts differed most for retention (+1.03), $t(473)=3.55$, $p=0.00$, purpose (0.93), $t(223)=5.05$, $p=0.00$, and sale (+0.82), $t(221)=4.54$, $p=0.00$.

In our sample, 59% of privacy experts and 49% of respondents indicated that they would rate the attributes differently for different types of services. This is in line with similar findings indicating that privacy concerns are dependent on the context and the type of service (K. Martin & Shilton, 2016; Nissenbaum, 2004; Nissenbaum, 2011; Phelps et al., 2000; Xu et al., 2008).

Previous research suggests that privacy concerns are influenced by demographic factors (Bellman et al., 2004; Zukowski & Brown, 2007). To investigate whether men and women felt differently about their privacy, we ran an independent sample t-test for all 15 privacy attributes. No significant differences were found, which is consistent with the results of a recent meta-study (Tifferet, 2019). Since age is often found to be associated with privacy expectations (Bellman et al., 2004), we ran an ANCOVA to control for the age of the respondents in assessing the differences between the mean scores of the users and the privacy experts. The only significant difference we found was for the right to be forgotten ($p = 0.01$), but the adjusted means were almost the same as the unadjusted means. Therefore we conclude that age is not a confounding variable. The vast majority of the respondents were European nationals. Since all Europeans fall under the same privacy regime, controlling for nationality was deemed unnecessary.

Table 6.1 Occurrence rates of privacy attributes in literature

Privacy attributes	Privacy visualizations		Privacy by Design guidelines	
	Count	Percentage	Count	Percentage
Accountability	3	33%	11	11%
Anonymization	3	33%	6	6%
Collection	11	111%	11	11%
Control	8	80%	12	12%
Correctness	4	40%	11	11%
Disclosure	4	40%	9	9%
Functionality	1	10%	1	1%
Pseudonymization	2	20%	7	7%
Purpose	12	120%	11	11%
Retention	8	80%	11	11%
Right to be forgotten	2	20%	4	4%
Sale	2	20%	0	0%
Security	6	60%	14	14%
Sharing	12	120%	7	7%
Transparency	7	70%	14	14%
	Total		Total	
	33	330%	111	111%
	Privacy Label (2020)		Privacy Company (2018)	
	DAPIS (2019)		OASIS (2016)	
	CleverFRANKE (2019)		Hoepman (2014)	
	Fox (2018)		OECD (2013)	
	DCIS (2012)		ISO29100 (2011)	
	GDPR draft icons (2014)		Cavoukian (2009)	
	Privacy wheel (2012)		GAP (2009)	
	TrustArc (2011)		ISTPA (2007)	
	PrimeLife (2010)		GFS (2006)	
	Mozilla (2010)		APFC (2005)	
	CyLab Nutrition label (2009)		CSA (2001)	
	KnowPrivacy (2009)		APPs (2001)	
	Mehldau (2007)			

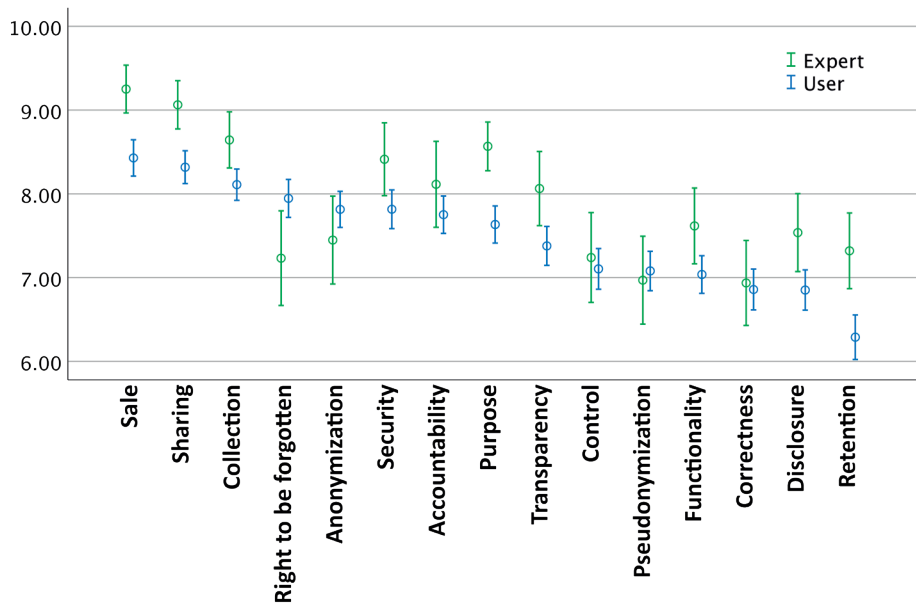


Figure 6.14 Mean importance (1-to-10) and confidence interval of privacy attributes, sorted by mean importance to users

6.5 Discussion

Our literature review (Table 6.1) revealed notable differences between privacy visualizations and PbD guidelines in terms of the privacy attributes they cover. Additionally, experts and users rated some attributes differently in the survey of Section 6.4.4. In this section, we examine similarities and differences between these four perspectives: Privacy visualizations, PbD guidelines, privacy experts, and users.

On average, PbD guidelines cover more attributes than visualizations (Mean = 8.6 vs 5.6 attributes per proposal). This result is not surprising if we consider that privacy visualizations are mostly designed to provide simple, user-friendly information about the handling of personal data (Metzger, 2007; Pangrazio & Selwyn, 2019). In our survey, privacy experts assigned a higher importance to most attributes, on average. This is to be expected because as privacy officers, they are not only concerned with privacy as users, but also professionally.

Sale and *sharing* were rated as the most important attributes by most users and privacy experts in our sample. However, while icons related to the sharing

of data were included in all but one of the privacy visualizations, only half of the PbD principles dealt with this issue. Studies find that willingness to exchange personal data is strongly mitigated by secondary use (K. Martin & Shilton, 2016; Solove et al., 2006) so it makes sense that almost all of the privacy visualizations we reviewed describe data sharing since they are aimed at users.

Sale—the attribute consistently ranked as most important in our survey—is covered by just two privacy visualizations and zero PbD guidelines. This is evidence of a growing discrepancy: while the sale of personal data remains an intrinsic part of the business model for online service providers (K. Martin, 2016b; Rothfeder, 1992) it is one of the major concerns of users (Kozyreva, et al., 2020).

Collection and *purpose* are arguably the most fundamental privacy attributes because they describe which data are to be collected and why. The privacy experts we surveyed consider both collection and purpose to be of very high importance (closely following sale and sharing). Other studies confirm this observation (Ackerman et al., 1999). Users in our sample, however, rate purpose as less important. We speculate this is because users consider certain types of data as sensitive regardless of purpose (Bansal et al., 2010). Nevertheless, collection and purpose were the most frequently occurring attributes in both privacy visualizations and PbD guidelines. Therefore, they appear to be the most important attributes to consider when discussing online privacy.

Transparency was mentioned in all PbD guidelines but only half of the privacy visualizations. We speculate this is because privacy visualizations are themselves a tool for achieving and showing transparency.

Security of personal information is mentioned by all of the 14 PbD guidelines we reviewed, but less than half of the visualizations, mostly those published after 2012. In our survey, privacy experts ranked security as the fifth most important attribute (users ranked it as sixth). This suggests that the security of personal information is considered critical for developing privacy-aware online services, but is also of increasing concern to users.

Accountability is also mentioned more often in proposals for PbD guidelines than for visualizations (almost 80% vs. 23%). This is not surprising, since accountability increases the magnitude of potential losses for the service provider in case of data breaches and PbD guidelines are aimed at developers. Nevertheless, accountability was ranked as the seventh most important attribute by users in our sample.

Retention is ranked significantly higher by privacy experts compared to users and also covered by most PbD guidelines and privacy visualizations. The right to be forgotten, on the other hand, was perceived as more important by

users and is rarely mentioned in the privacy visualizations or PbD guidelines we reviewed. The right to be forgotten and retention both relate to the ability of an organization to delete privacy sensitive data. However, most IT systems that are in operation today have not been designed to allow the data controller to delete data from all possible sources in which the data may reside. While retention has always been a consideration when it comes to data handling, the right to be forgotten is a relatively new, user-driven initiative. This is supported by the fact that in our literature review, we only found one mention of it before 2011. Indeed, managing legacy data sources in a GDPR-compliant manner is a major challenge to data controllers (Perera et al., 2016) and privacy experts are well aware of this. Knowing how hard it is to completely remove data from all sources might cause our privacy experts to rate the importance of the right to be forgotten lower than the users we surveyed. It is quite possible that the average user has unrealistic expectations regarding the ability of organizations to erase all data items pertaining to them.

Anonymization was ranked as the fifth most important attribute by users and the eighth most important attribute by privacy experts, but received surprisingly little attention in literature. Anonymization is technically challenging (Morales-Trujillo et al., 2018) and privacy experts know this. Because true anonymization is seldom achievable (Pedarsani & Grossglauser, 2011; Wondracek et al., 2010), various degrees of pseudonymity are implemented instead. Although the information security practitioners we interviewed felt that pseudonymization should be differentiated from anonymization, several privacy experts in our survey indicated that the two attributes are difficult to distinguish. We speculate users are even less familiar with this distinction, which explains why they ranked pseudonymization as one of the least important attributes. Nevertheless, taking steps to remove personal identifiers from user data is of interest to users, which also implies this should be given more careful consideration by developers. However, from a practical perspective, pseudonymization can be viewed as partial or imperfect anonymization.

Control and *correctness* were ranked relatively low by users as well as privacy experts but were often encountered in PbD guidelines. Furthermore, correctness was represented in 30% of visualizations. Online services increasingly gather and aggregate user data to glean insights into habits, trends or behaviors not directly related to the actual exchange of the product or service (A. L. Allen, 2016; K. Martin, 2016b; Myers West, 2019), but privacy controls are widely perceived as overly complex by users (Beznosov et al., 2009; Ramokapane et al., 2019). The resulting difficulty in managing personal data results in privacy fatigue: a sense of not being in control of the collection and sharing of data

online (Choi et al., 2018; Hoffmann et al., 2016). This weakens the perceived utility and therefore importance of privacy settings and controls. Nevertheless, such mechanisms enhance privacy both proactively (preventing unauthorized collection or collection of incorrect data) and reactively (consent withdrawal and the correction of previously collected data). Therefore, providing control over data collection and maintaining correctness of user data is an inherent part of online privacy (K. Martin, 2016b).

6.5.1 Trends

Although we reviewed PbD guidelines published or updated after 2001, our initial search returned many older PbD guidelines. The FTC Fair Information Practice (FIPPS) was the first set of PbD principles, forming the foundation for many of the newer principles and legislation (U.S. Department of Health, Education and Welfare, 1973; The Privacy Act of 1974). In 1990, the UN published similar guidelines (UN General Assembly, 1990) and in 1995, the EU introduced its first Data Protection Directive (The European Parliament and the Council of European Union, 1995).

Throughout the first decades of the 21st century, the publication rate of PbD guidelines slowly increased and after 2009 we saw an increase in domain- or technology-specific PbD guidelines. Since most of the PbD guidelines we found are either regulation or industry standards, we conclude that privacy by design has made its way into practice.

On the other hand, all of the privacy visualizations we found were published after 2007, with the majority being published by academics after 2012. This coincides with an increase in privacy awareness. Although the need for communicating online privacy is not a new discussion (Metzger, 2007), research into empowering users to make informed disclosure decisions has recently started to gather steam (L. Edwards & Abel, 2014; Holtz et al., 2011a; Pangrazio & Selwyn, 2019; Rossi & Palmirani, 2017). We are starting to see industry initiatives as well. However, despite the fact that both the European GDPR (The European Parliament and the Council of European Union, 2016) and the US Federal Trade Commission (FTC; Anthony, 2001) recommend standardized privacy labels, no official standard has yet been defined.

Disclosure, correctness, accountability, and the right to be forgotten are increasingly common in recent privacy visualizations. This trend likely reflects increasing concerns regarding safe harbor (Colonna, 2013; The European Parliament and the Council of European Union, 2000) and data breaches (B. Edwards et al., 2016). Even though correctness and accountability are covered by many PbD guidelines, disclosure is not covered by recent initiatives such

as PIPEDA (Office of the Privacy Commissioner of Canada, 2000), the Privacy Commission (2019), and Privacy Label (2020).

Sale, the right to be forgotten, anonymization and accountability were rated as very important by our sample of users. However, accountability and anonymization are missing from most privacy visualizations while sale and the right to be forgotten are missing from PbD guidelines as well. But sale of personal data is of increasing concern to users, EU law mandates the right to be forgotten, anonymization is becoming an industry standard, and service providers have been receiving record fines for privacy infringements. These developments lead us to believe that, while current approaches to communicating and implementing privacy do not yet take the needs and preferences of users into account, this situation will (hopefully) change in the future.

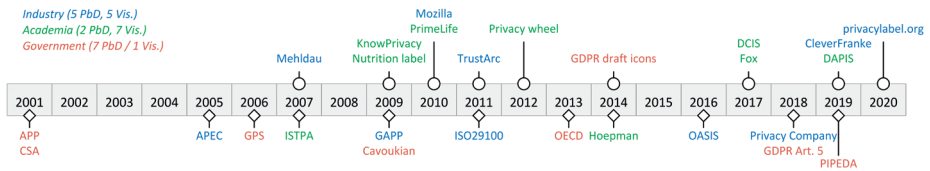


Figure 6.15 Publication timeline of privacy visualizations and PbD guidelines

6.5.2 Limitations

Because the entry point of the literature search was Scopus, it is possible that not all relevant proposals from industry were considered. We mitigated this by performing auxiliary Google searches. Furthermore, even though we ran several searches using nine synonyms for principles and fourteen synonyms for visualizations, important keywords may have been missed. We do believe however, that our sample of 27 proposals is sufficient to reach saturation in terms of privacy attributes. This is supported by the fact that each privacy attribute was encountered in at least two documents and that over 93% of privacy experts and users we surveyed indicated the unified list was complete and unambiguous.

Some of the documents selected for our systematic review were ambiguous and many differed in terms of granularity and scope. Therefore, multiple attributes were sometimes attached to the same principle or visualization and multiple principles or visualizations sometimes corresponded to a single privacy attribute. Nevertheless, after three rounds of coding we reached almost perfect agreement between coders. This indicates that, while some of the attributes in the list may be grouped together, the list itself is complete and understandable.

In our online survey, the expert sample was smaller than the user sample. This is because privacy experts are a specialized group and a larger sample was hard to obtain. A disproportionate number of the respondents were young and have attended higher education. However, age and gender were not found to be confounding variables. While the list of attributes is international, almost all respondents were European which makes our ranking European.

The results might be influenced by response bias. However, the topic of our questionnaire is not socially sensitive and therefore the risk of giving socially desirable answers is small. Furthermore, by screening the raw data rigorously and removing superficial and incomplete responses, we are confident that we have managed to keep any potential response bias to a minimum.

Lastly, differences between the perceived importance of most attributes were small and many respondents indicated that their rating depends on the type of application. We mitigated this by also considering the occurrence rate of each attribute in the literature we reviewed.

6.5.3 Practical recommendations

6.5.3.1 Privacy visualizations should be legally mandated

Except for CLEVER°FRANKE's (Franke et al., 2019), DAPIS (Rossi & Palmirani, 2019) and Privacy Label (2020)—which are currently under development—all of the other privacy visualization projects have been abandoned. We speculate that adopting such labels—and more importantly, getting a good score—provides a non-functional benefit to the user but comes at great costs for the provider, as is often the case with safety and security. Indeed, third-party privacy seals are not correlated with trustworthiness (Edelman, 2009) and crowdsourcing efforts such as TOS:DR⁹ have so far been unsuccessful. Providers should therefore supply an understandable summary of their privacy policies themselves (H. Wang et al., 1998). However, since similar endeavors such as the EU energy label, movie ratings, and even seatbelts had to become mandatory before they were adopted, privacy visualizations will only become wide-spread if they are legally mandated.

9 <https://tosdr.org/>; Terms of Service; Didn't Read is a project inspired by Aza Raskin's Privacy Icons (2010) and EFF's TOSBack (<https://tosback.org/>; a project of the Electronic Frontier Foundation, eff.org)

6.5.3.2 Privacy visualizations should go beyond data collection and processing

We find that most privacy labels align with Nissenbaum (2011) and K. Martin and Shilton (2016) in that they primarily communicate what information is collected, how this information is shared and for what purpose. However, our ranking suggests that sale of data must also be made explicit. Furthermore, although most current visualizations do not include an indication of the level of security and accountability, this is important to both privacy experts and users and actually mandated by the GDPR (The European Parliament and the Council of European Union, 2016). Trustworthy online data exchange relies on obtaining truly informed consent (J. Martin & Christin, 2016), and this requires providing the end-user with the relevant information in an understandable form. Our ranked list of privacy attributes serves as a basis for a user-centric privacy visualization which covers all of the important aspects of privacy.

6.5.3.3 PbD guidelines should be more user-centric

One of the most striking findings was the fact that the two attributes rated as most important by both privacy experts and users (sale and sharing) were rarely covered by PbD principles. To avoid anxiety, uncertainty, or even fear (Nissenbaum, 2011), the gap between privacy concerns and guidelines aimed at addressing them must be reduced. PbD is aimed at taking the privacy concerns of the end-user into consideration during development, and so issues related to data sharing, and in particular sale of user data must be part of PbD guidelines. Ideally, since the lowest average importance rating was six on a scale going from 1 to 10, PbD guidelines should cover all of the attributes on our list, with the possible exception of functionality. This is because functionality was ranked as one of the least important attributes and was sometimes marked as confusing by both privacy experts and users.

6.5.3.4 The right to be forgotten should not be forgotten

The right to be forgotten was rarely mentioned in the PbD guidelines we reviewed. In 2014 however, the European Court of Justice ruled that European users can request the removal of personal data from online service providers (González v. Google Spain, 2014) and the GDPR mandates this as well (despite the fact that the right to be forgotten is not one of the GDPR's PbD principles). Newman (2015) questions the extent to which the right to be forgotten is financially and legally feasible. Still, according to Ausloos (2012), the ability to demand the erasure of personal data can and must be available in data processing situations where consent was required and—with normative, economical, technical, and legislative changes—this could be implemented more widely. Even though most PbD

guidelines already recommend obtaining consent (i.e. control) and recommend removal of data when it is no longer necessary (i.e. retention), the right to be forgotten goes a step further by giving users the ability to withdraw consent. Therefore, the right to be forgotten (or its diluted form, the “right to erasure”; Ambrose & Ausloos, 2013), should be an integral part of future PbD guidelines.

6.5.4 Research challenges

6.5.4.1 Structuring privacy policies

Privacy policies often focus on collection, sale, and sharing of user data, but our survey revealed that the right to be forgotten and security are of increasing concern. Furthermore, regulation increasingly mandates that privacy policies provide information about potential disclosure to (foreign) government entities, accountability in case of breaches, and the ability to correct one’s data. The unified list of privacy attributes of Section 6.4.3 is based on extensive review and comparison of privacy attributes covered by privacy visualizations and PbD guidelines aimed at online services in general. Therefore, it represents a complete and technology-/domain-independent checklist of aspects related to online privacy. A valuable research direction is to investigate whether such a checklist can be used to verify the completeness of privacy policies (Al-Jamal & Abu-Shanab, 2015) or to structure – or even automatically restructure – privacy policies (Yu et al., 2015).

6.5.4.2 Developing a privacy rating system

Similar to PrivOnto (Oltramari et al., 2018), the privacy attributes on our list can be operationalized so that they can be used measure and compare the privacy level of online services on multiple metrics. Such a rating mechanism can be used to classify online services based on their privacy policy and—in the long term—could provide a standardized, understandable, machine-readable summary of privacy policies that enables both providers and users to assess, communicate, and compare the privacy of online services. To explore this direction, we have started developing a free online service which implements some of these ideas: privacyrating.info. However, developing a usable and useful privacy rating poses a significant research challenge.

6.5.4.3 Investigating context dependency of privacy attributes

The unified list of privacy attributes in Section 6.4.3 is a first step toward a standardized list of privacy attributes that can function as the foundation of a privacy visualization. However, the work of Nissenbaum (2011) and

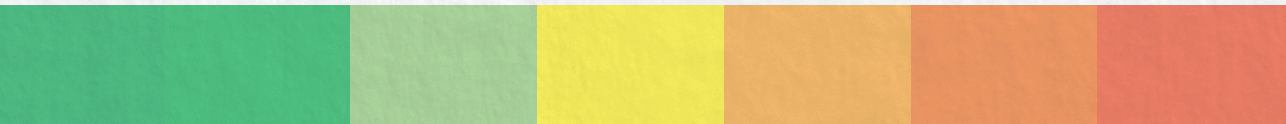
K. Martin (2016b) showed that information privacy is discriminate, embedded in the context, and based on a social contract between the various stakeholders involved in the information exchange. It would seem that privacy perception is not universal, but context-dependent (Nissenbaum, 2004; Phelps et al., 2000; Xu et al., 2008). However, Solove (2002) assumes that there is also a certain congruity between situations of personal data disclosure online. The extent to which privacy is context-dependent is an open problem. Is a universal privacy visualization effective? Is there a need to specialize? And if so, how might a tailored solution look like?

6.6 Conclusions

We performed a systematic review of current approaches to communicating privacy issues to users (privacy visualizations) and to developers (PbD guidelines). It revealed significant gaps in terms of the aspects of data processing these approaches cover. To understand these differences, we distilled a unified list of privacy attributes and ranked it based on perceived importance by European privacy experts and users.

Our study revealed that some attributes are considered important by both privacy experts and users: what type of personal data are collected, with whom it is shared with, and whether or not it is sold. The PbD guidelines we reviewed also emphasize collection, but mention purpose more often than sharing or sale. Furthermore, PbD guidelines often focus on ensuring information security and transparency while providing users with privacy controls. Privacy visualizations take a user-centric perspective, focusing on collection, purpose, and sharing. Overall, we see an increase in publications pertaining to PbD and privacy visualizations. The right to be forgotten and accountability of service providers are increasingly mentioned in both regulations and guidelines. Both were found to be important in our survey. Disclosure to law enforcement, retention periods, and correctness of data are also mentioned increasingly often in publication covering online privacy, although these were ranked as relatively unimportant by our sample of privacy experts and users. Pseudonymization, anonymization, and the trade-off between functionality and privacy are mentioned in a minority of the literature we reviewed and were perceived to be relatively unimportant by the users and privacy experts we surveyed.

The results serve as (1) a ranked list of privacy best-practices for developers and providers of online services, as (2) a foundation to visually communicate the most relevant aspects of a privacy policy to users, and (3) a taxonomy for structuring, comparing, and – in the future – rating privacy policies of online services.



7

Privacy Rating: A user-centered approach for visualizing data handling practices of online services

Barth, S., Ionita, D., De Jong, M. D. T., Hartel, P. H., & Junger, M. (in press). Privacy Rating: A user-centered approach for visualizing data handling practices of online services. *IEEE Transactions on Professional Communication*.

7.1 Background

Imagine giving a complete stranger your address and phone number, the contact information of everyone you know, unlimited access to your photos, a detailed account of your media use, all your private messages and real-time updates on your whereabouts. It sounds extreme, but most of us risk doing it every day—simply by using online services. Online services ranging from social media and entertainment, to shopping and banking continuously handle large amounts of our personal data. The pervasiveness of digital media in modern life has resulted in a ‘semantic web’ built almost entirely on personal data (Berners-Lee et al., 2001).

Using online services inevitably requires making decisions about disclosing personal data. Disclosures may have adverse consequences such as misuse, spam, or identity theft (Huckvale et al., 2015; Matz et al., 2020; Milne et al., 2004). However, due to the complex, multifaceted and intangible nature of online privacy, the vast majority of users have difficulty judging potential privacy risks and safeguarding their privacy (Alsaleh et al., 2017; Mourey & Waldman, 2020). Privacy policies detail how online services handle user data, but few users attempt reading them and those who do, face difficulties understanding them (Jensen & Potts, 2004; Prichard & Mentzer, 2017; Proctor et al., 2008; Rudolph et al., 2018; Sunyaev et al., 2015). Furthermore, an analysis of privacy statements showed that many such disclaimers place little emphasis on providing users with clear cut information designed to aid the decision-making process. In fact, self-interest and litigation avoidance have much higher priorities among most online service providers (Papacharissi & Fernback, 2005).

The complex, multifaceted and intangible nature of online privacy may amplify the cognitive biases users already have, including optimism bias (underestimating the risks of unsafe behaviors), status quo bias (exhibiting an affinity for default choices), app desirability bias (adjusting privacy concerns based on the desirability of the app) and anchoring (taking other users’ behaviors as a reference point) (Acquisti et al., 2017; Gu et al., 2017). A recent study showed that, in line with Festinger’s cognitive dissonance theory (Festinger, 1957), users tend to consider privacy less important when they think they are not in control anyway (Mourey & Waldman, 2020).

Online privacy does not occupy a prominent position on the research agenda in technical and professional communication, with only very few research articles in the last fifteen years devoted to the topic (Chai et al., 2009; S. Young, 2021)—none of which address the challenge of empowering users to act in accordance with their own privacy interests. We believe online privacy deserves more

attention within our discipline, as it is an increasingly prominent and inherently complex aspect of the interaction between humans and technology, which could benefit from the verbal and visual communication competencies that typically define the strength of our discipline.

When it comes to empowering users to assume informed responsibility for their online privacy, many researchers have drawn attention to the potential of using privacy labels, visually depicting the privacy threats associated with online services (S. De Jong & Spagnuolo, 2020; L. Edwards & Abel, 2014; Efroni et al., 2019; Esayas et al., 2016; Fox et al., 2018; Holtz et al., 2011a, 2011b; Kelley et al., 2009, 2010; Petterson, 2015; Renaud & Shepherd, 2018; Rossi & Palmirani, 2017, 2019; Soumelidou & Tsohou, 2019; Tesfay et al., 2018; Van den Berg & Van der Hof, 2012; Van Kleek et al., 2017). In fact, the European GDPR mandates ‘standardized icons’ as an overview of the intended data processing (The European Parliament and the Council of European Union, 2016). In this article, we describe the development and evaluation of *Privacy Rating*, a new privacy visualization we have developed for online services. The label is the result of a research-based inventory of important privacy risks, includes an efficient tool for mapping privacy features and has a design aimed at raising privacy awareness among non-engaged users and providing relevant information to users who already have higher levels of privacy concerns. After a literature review, we describe the privacy label and its rationale before reporting on the design and the results of a user test that focused on usability, perceived usefulness and the effects on users’ trust in an online service.

7.2 Literature review

7.2.1 Why is there a need to visualize privacy?

Although users claim to care about their online privacy and have concerns about privacy violations, they generally do not behave accordingly: They download apps, give permissions and provide personal information without giving the potential ramifications of their actions much thought. This attitude-behavior discrepancy is known as the ‘privacy paradox’ (Barnes, 2006). Research shows that there may be three underlying mechanisms: (1) simply put, users rationally weigh the benefits of downloading an app, giving permissions or providing personal information against the associated privacy risks, (2) users have trouble weighing costs and benefits and instead rely on (possibly biased) heuristics or cognitive shortcuts, and (3) users do not even consider the privacy aspects of downloading an app, giving permissions, or providing certain information

(Barth & De Jong, 2017). The distinction between these mechanisms may in practice not always be clear. Through their behavior, users put themselves at unnecessary risk. The current situation is a vicious circle: Virtually all privacy policies are complex and ‘take it or leave it’, therefore, individual users have no real choice but to accept online services on their (unclear) terms, a situation that panders to the strategies of many service providers. Although online privacy is a topic of vivid discussions in the academic literature, in practice it is often reduced to momentary feelings of unease and uncertainty in users.

Designers and providers of online services are in the best position to make data handling processes more transparent to users. Since the end of last century—long before the introduction of smartphones—researchers have advocated and worked on a Privacy by Design paradigm (Cavoukian, 2011b; Langheinrich, 2001). The basic premise is that privacy should be incorporated into the fabric of online services instead of bolting it on after the fact. Many Privacy by Design standards and guidelines have emerged, for instance by ISO (Technical Committee ISO/IEC JTC 1SC 27, 2011). Although this approach can make a tremendous contribution to users’ online privacy, several authors have warned of the legal and practical complications (Klitou, 2012; Koops & Leenes, 2014), as well as the problems of adoption and implementation (Bu et al., 2020; Cavoukian, 2020; Gerunov, 2020). In practice, many providers of online services still try to discourage users from exercising their rights to privacy (Forbrukerrådet, 2018). Additionally, a core characteristic of online services is personalization, which by definition involves some degree of personal data processing. Research shows that different users may have different tolerances to specific data handling practices (Barth, Ngo, et al., 2020).

Empowering users consciously take more responsibility for their online privacy would be another solution. This could entail increased education: providing users with, for instance, more knowledge about the business models of online services, the potential privacy risks of transactions, the exact meanings of permissions and the best protection methods. However, research suggests that general knowledge and privacy awareness play no significant role in the privacy paradox: Advanced Computer Science students and even privacy and security experts appear to struggle with the same issues as lay users, exhibiting similar unsafe behaviors (Barth et al., 2019; Barth, De Jong, et al., 2020).

From a document design perspective, there may be a lot to gain from better provision of privacy risk information. Given the shortcomings of current privacy statements (Jensen & Potts, 2004; Papacharissi & Fernback, 2005; Prichard & Mentzner, 2017; Proctor et al., 2008; Rudolph et al., 2018; Sunyaev et al., 2015), some researchers investigated whether or not textual

improvements could help. An experimental study showed that merely simplifying privacy statements based on document design principles will not affect users' comprehension, attitudes, or behavior (Ben-Shahar & Chilton, 2016). On the other hand, another experimental study showed that concise and simple privacy warnings do have an effect on users' risk perceptions and online behavior (LaRose & Rifon, 2007). Beyond their legal jargon, and word-, sentence- and paragraph-level complexity—all severe problems in their own right—privacy statements generally represent an intimidating information overload that does little to align with the perspective of users trying to ascertain whether to use an online service or not. It seems important to realize that there is functional complexity involved when communicating privacy risks (M. D. T. De Jong & Wu, 2018; Lentz & Pander Maat, 2004). Ideally the same privacy information should:

- Raise users' awareness of the importance of privacy and privacy risks (Deuker, 2010; Pöttsch, 2009);
- Provide less engaged users with a shortcut to support their decision-making as it pertains the potential privacy risks associated with using an online service;
- Provide highly engaged users with user-friendly and comparable information about privacy risks (with varying levels of detail, depending on their interests and expertise).

Privacy visualizations, as advocated and developed by several researchers (S. De Jong & Spagnuolo, 2020; L. Edwards & Abel, 2014; Efroni et al., 2019; Esayas et al., 2016; Fox et al., 2018; Holtz et al., 2011a, 2011b; Kelley et al., 2009, 2010; Petterson, 2015; Renaud & Shepherd, 2018; Rossi & Palmirani, 2017, 2019; Soumelidou & Tsohou, 2019; Tesfay et al., 2018; Van den Berg & Van der Hof, 2012; Van Kleek et al., 2017), may be a viable way to address the communication challenge. More than verbal information, they can draw the attention of users who are not aware of privacy risks (L. Edwards & Abel, 2014; Sheng et al., 2020; Soumelidou & Tsohou, 2019) and force the designer to translate complex privacy information into manageable, standardized privacy information.

7.2.2 Earlier attempts to visualize online privacy

Developing a privacy visualization requires two related activities: (1) an intrinsic analysis of the relevant privacy aspects to be included, and (2) a verbal-visual communication design. Both in the academic literature and in practice, many attempts have been made to develop privacy visualizations (see Barth, Ionita, et al., 2020). Table 7.1 summarizes 14 earlier attempts, with special attention on

the extent to which the systems provide overall advice about the privacy risks of online services (overall indicator) and detailed information about specific privacy aspects (privacy details).

Existing privacy visualizations operationalize privacy quite differently (Barth, Ionita, et al., 2020; S. De Jong & Spagnuolo, 2020). Barth, Ionita, et al. (2020) investigated operationalizations of online privacy that manifest themselves in Privacy by Design guidelines and privacy visualizations, resulting in 15 different privacy aspects. Not one of these privacy aspects was incorporated in all privacy visualizations. Three aspects were quite prominent—types of data collection, purposes of data collection, and data sharing—with only one or two visualizations missing out, but the overall focus of the visualizations differed considerably. An agreed-upon framework of relevant privacy aspects as they pertain to online services does not currently exist. A new privacy visualization should thus be based on a systematic analysis of relevant aspects of online privacy.

Various types of visualizations can be distinguished. Seven of the 14 visualizations listed in Table 7.1 are icons sets expressing specific privacy characteristics. Several authors mentioned that it is hard to visualize such intangible and complex features (Esayas et al., 2016; Hansen, 2009; Van den Berg & Van der Hof, 2012), and several icons that were developed indeed proved to be problematic in user tests (Graf et al., 2011; Holtz et al., 2011a). Likely because of that, some of the icon sets use supporting tags to assist visual cue interpretation. A significant drawback of icons is that they are limited to depicting specific privacy risks—making them unsuitable for providing users with ‘the bigger picture’ that is necessary if they are to make informed decisions about the acceptability of the combined privacy risks.

Three of the proposed visualizations downplay the role of icons by making them merely supportive for predominantly written information. In these cases, the icons have no independent meaning, but only visually support the structure of a summarized privacy text. Again, it is questionable whether or not this approach supports users in their decisions about the combined privacy risks of online services. The difficult task of making sense of the various privacy characteristics is still entirely the users’ responsibility.

Table 7.1 Overview of earlier privacy visualizations

Year	Source	Name	Type	Overall Indicator	Privacy Details
2007	Mehldau (2007)	Mehldau's data privacy declarations	Icons with tags	No	Yes
2009	Kelley et al. (2009, 2010)	CyLab's privacy nutrition label	Table	No	Yes
2009	Gomez et al. (2009)	KnowPrivacy's policy coding	Icons with tags	No	Yes
2010	Moskowitz and Raskin (2011)	Mozilla's privacy icons	Icons	No	Yes
2010	Graf et al. (2011); Holtz et al. (2011a, 2011b)	PrimeLife privacy icons	Icons	No	Yes
2011	Pinnick (2011)	TrustArc's privacy short notice	Icons	No	Yes
2012	Van den Berg and Van der Hof (2012)	Privacy wheel	Privacy label	Yes	Yes
2014	Petterson (2015); Esayas (2016)	GDPR's draft privacy icons	Icons	No	Yes
2017	Van Kleek et al. (2017)	Data controller indicators	Data flow representation	No	Yes
2018	Renaud and Shepherd (2018)	Renaud and Shepherd's privacy summary	Summary supported by icons	No	Yes
2018	Fox et al. (2018)	Fox et al.'s GDPR compliant privacy label	Summary supported by icons	No	Yes
2019	Rossi and Palmirani (2017, 2019)	Data protection icon set (DaPIS)	Icons with tags	No	Yes
2019	Franke et al. (2019)	Clever°Franke's privacy label	Privacy label	Yes	Yes
2020	Privacy Label (2020)	Privacy label	Summary supported by icons	No	Yes

Two other visualization proposals explore very different directions. Inspired by the nutrition labels on food, Kelley and colleagues developed a privacy nutrition table, which actually consists of a listing of ten types of user data, five types of data handling, and two different parties handling the data. In each cell of the table, four options may be entered (yes, no, opt out, and opt in) (Kelley et al., 2009, 2010). The underlying metaphor of nutrition labels suggests that the visualization does not focus on less-engaged users and does not support users' overall decisions on acceptable or non-acceptable privacy risks. Still, a focus group study showed that users appreciated the system (Kelley et al., 2009) and a comprehensive experiment showed that the label, compared to normal privacy statements, helped users better understand the privacy aspects of online services (Kelley et al., 2010). Van Kleek and colleagues developed a visualization of the data flows from online services (Van Kleek et al., 2017). Although the resulting graphs were advanced and may be too complex to be intuitively comprehensible, a small-scale experimental study indicated that the visualization, more than written privacy information, helped users when making informed decisions regarding online privacy.

Finally, two proposals for visualizations take the form of privacy ratings, providing overall indications of the privacy aspects of online services with optional in-depth information. Van den Berg and Van der Hof's privacy wheel (Figure 7.1) consists of an overall privacy qualification in the middle encompassed by eight brightly colored clickable privacy aspects (Van den Berg & Van der Hof, 2012). Although it manages to combine an overall privacy assessment and provide more detailed information, the visualization has a few potential drawbacks: The overall privacy assessment in the middle might be easily overlooked, lacks a reference point and is not transparently related to the eight specific privacy aspects. Clever^oFranke's privacy label (Figure 7.2) is inspired by the letter classification (A-F) and color use of the EU energy label. It consists of a colored circle with a privacy qualification in the middle: An A (in green) is positive, an F (in red) is negative (Franke et al., 2019). Around the qualification, there is a circle divided into three equal parts representing three privacy aspects: data usage, data collection, and user control. For every aspect, five questions are asked. In the case of a positive answer the line is colored; in the case of a negative answer it is left white. The thicker the colored circle around the privacy qualification, the more positive the service scores on the specific privacy aspects. Users can use a QR code for more specific information. A drawback of this visualization is that the specific privacy information is hidden in the design and that the system of five questions in three parts of the circle may not be clear to users. There are no research reports available on user tests with either privacy label.



Figure 7.1 Van den Berg and Van der Hof's Privacy Wheel (2012)



Figure 7.2 Clever°Franke's Privacy Label (Franke et al., 2019)

In this article, we describe a project developing a privacy rating tool for online services that is founded upon expert knowledge on the relevant privacy aspects and designed to overcome the perceived shortcomings of the earlier privacy visualizations. Furthermore, we describe a user study focusing on usability, perceived usefulness and any effects on user trust. This leads to the following research questions:

RQ1.How can we design a privacy rating tool that optimally empowers users with different levels of privacy knowledge and awareness?

RQ2.How do users react to such a privacy rating tool, in terms of usability, perceived usefulness and trust in an online service?

7.3 The Privacy Rating

Below we will describe the development of the *Privacy Rating* tool. We will discuss its three main characteristics: content, visual design and implementation.

7.3.1 Content: Privacy aspects and their operationalization

The development of *Privacy Rating* started with a thorough and systematic analysis of the privacy aspects of online services that should be deemed relevant and therefore included. We took the list of 15 privacy attributes gathered in earlier research (Barth, Ionita, et al., 2020) as our starting point (Table 7.2). The attributes were based on established Privacy by Design guidelines and earlier privacy visualizations. Research with experts and users confirmed the importance of all attributes (Barth, Ionita, et al., 2020). For intrinsic reasons, we decided to exclude two of the original attributes for our visualization. Functionality was removed because it was too ambiguous and partly covered by control. Transparency was removed because having a privacy rating can already be seen as a positive indicator of transparency in itself. In addition, anonymization and pseudonymization were collapsed into one attribute, as they were sometimes hard to distinguish: Pseudonymization can be seen as incomplete anonymization. From previous research, we know that privacy is subjective and context-dependent (Barth, Ngo, et al., 2020; Nissenbaum, 2011, 2019). Therefore, we decided to use all of the remaining 12 attributes as equal metrics for our rating system.

As twelve different privacy attributes are not manageable for users, we conducted a card-sort study in which we asked users to cluster the 12 attributes. Most often, the attributes were grouped into four categories. While security turned out to be a clear group label, there was disagreement about the others. Consulting 10 privacy and cybersecurity experts from our network resulted in four main clusters: *Collection*, *Sharing*, *Control*, and *Security* (see Table 7.3).

Table 7.2 Privacy aspects considered for the Privacy Rating (Barth, Ionita, et al., 2020)

Privacy Aspect	Description
Accountability	Can the service provider be held accountable for violations?
Anonymization**	Are all identifiable markers completely removed so that data can never be traced back to individuals users?
Collection	Which user data are collected?
Control	Are users able to choose or decide which data to share for which purpose, and how difficult is it to do so?
Correctness	Are there mechanisms for preventing and fixing incorrect data?
Disclosure	What is the provider's attitude toward data requests from law enforcement?
Functionality*	Are users forced to choose between functionality and privacy?
Pseudonymization**	Are personally identifiable markers replaced by artificial identifiers, or pseudonyms, so that data can only be traced back to individual users with the help of additional information?
Purpose	What are the collected data used for?
Retention	How long are collected data stored?
Right to be forgotten	Can users request that all their personal data will be removed?
Sale	Are any of the data sold to third parties?
Security	Which technical measures are taken to ensure that data are protected from unauthorized or malicious access?
Sharing	Do any of the collected data leave the ownership of the provider?
Transparency*	Are users able to obtain information about how their personal data are handled?

Note. * = Removed from the *Privacy Rating* attributes; ** = Collapsed into one attribute.

To use the metrics for rating and comparing online services, they must be operationalized. To keep the system simple and understandable for users, we defined three-point scales (good-neutral-bad) for each attribute. In iterative sessions with privacy and cybersecurity experts, we arrived at the operationalized metrics presented in Table 7.3. Online services receive penalty points depending on their score on each metric (0 points for good scores, 1 point for neutral scores, and 2 points for bad scores).

The total number of penalty points is then used to categorize online services into seven classes, from A (lowest privacy risks) to G (highest privacy risks):

- Class A: 0 or 1 points
- Class B: 2 to 5 points
- Class C: 6 to 9 points
- Class D: 10 to 13 points
- Class E: 14 to 17 points
- Class F: 18 to 21 points
- Class G: 22 to 24 points

Table 7.3 Clustered and operationalized privacy attributes

Cluster	Attribute	Operationalization
Collection	Collection	0 - Collects anonymous data 1 - Collects personal data, relating to an identified or identifiable person 2 - Collects sensitive data, involving racial or ethnic origin, political views, religious or philosophical beliefs, trade union membership, genetic or biometric data, health status, or sexuality and sexual orientation
	Purpose	0 - Used for functionality only 1 - Used for customization (personalization in the current interaction) 2 - Used for profiling
	Retention	0 - Data not stored 1 - Data stored for a pre-determined limited time 2 - Data stored indefinitely
Sharing	Sharing	0 - No sharing of user data 1 - Sharing of anonymous user data 2 - Sharing of user data
	Sale	0 - No sale of user data 1 - Sale of anonymous user data 2 - Sale of user data
	Disclosure	0 - Statutory disclosure to local law enforcement (inside user's jurisdiction) 1 - Disclosure to local law enforcement (outside user's jurisdiction) 2 - jurisdiction) Disclosure to foreign law enforcement
	Control	0 - Opt-in (users must explicitly opt-in to allow data collection) 1 - Opt-out (data are collected by default, but users can opt-out) 2 - No opt-in or opt-out
	Right to be forgotten	0 - Data deleted upon request 1 - Data hidden upon request 2 - Data cannot be removed

Table 7.3 Continued.

Cluster	Attribute	Operationalization
	Correctness	<ul style="list-style-type: none"> 0 - All data can be amended 1 - Some data can be amended 2 - Data cannot be amended
Security	Security	<ul style="list-style-type: none"> 0 - Industry standard security (certified compliant with the latest version of either ISO 27001 or NIST 800-53) 1 - Basic security (developed in compliance with the OWASP Top 10 standard and tested according to the OWASP Application Security Verification standard or the OWASP Mobile/Web Security Testing Guide or equivalent) 2 - None of the above
	Anonymization	<ul style="list-style-type: none"> 0 - Anonymous (all identifiable markers are completely removed so that collected data can never be traced back to individuals) 1 - Partially anonymous (personally identifiable information fields within collected data are replaced by artificial identifiers or pseudonyms, so that data can only be traced back to individuals with additional information) 2 - Not anonymous (personally identifiable information is stored)
	Accountability	<ul style="list-style-type: none"> 0 - Legally accountable 1 - Legally binding privacy policy 2 - Not legally accountable

Note. 0-2 represents the number of penalty points for the alternatives.

7.3.2 Visual design

Our *Privacy Rating* (see Figure 7.3) was designed through an iterative process in collaboration with a professional design agency. Simplicity, clarity, recognizability, and attractiveness were important criteria throughout the design process. With its stable and marked overall design, the visualization has the potential to draw attention to privacy issues across different online services. The use of overall privacy classes helps less-engaged users to make a quick overall judgment about the potential privacy threats of online services. Similar to the familiar energy label, privacy classes are indicated by combinations of letters and colors (ranging from A plus green for the most positive online services; to G plus red for the most negative ones). The colors also reflect the conventional color scheme of traffic lights. The presence of a full scale helps users to interpret the score of a particular online service.

Users who are more engaged with online privacy are helped with two levels of more specific information. The first level, immediately visible in the visualization, involves the scores of the online service in the four main categories of privacy aspects (Collection, Sharing, Control, and Security),

which can have different colors depending on the specific scores. Each category is listed with its name and an icon. The second level, which can be reached by hovering over or clicking the main categories, provides more detailed information about specific privacy aspects.

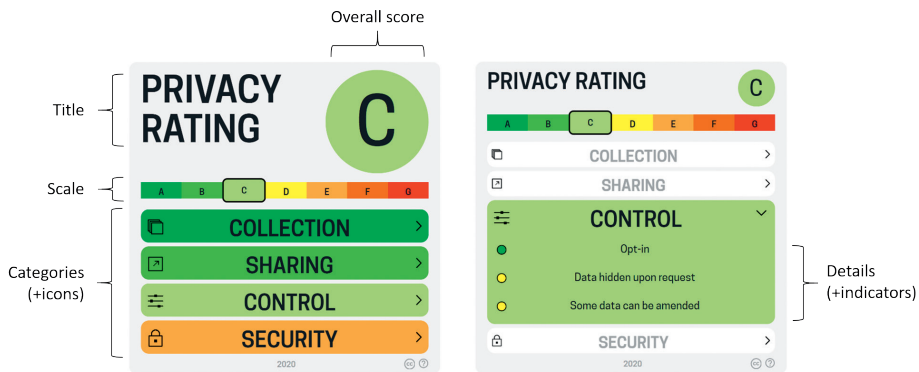


Figure 7.3 Design of the Privacy Rating

7.3.3 Implementation

To promote the practical feasibility of the *Privacy Rating*, we developed a self-assessment form in a free web application (www.privacyrating.info). The application is designed to walk service providers through a questionnaire with each question corresponding to one of the three levels of each attribute. The questionnaire is interactive: Once a question confirms the level of an attribute, the remaining questions corresponding to that attribute are skipped and the service provider is directed to questions about the next attribute. When all 12 attributes are evaluated, the application computes the privacy rating and creates a visualization in two formats: a HTML and a smaller PNG version, both of which can be embedded into web pages or apps. The small version can be added to the footer of the page or to the cookie notice and the larger version can be included in the privacy policy or as a pop-up.

7.4 Research design of the user study

To evaluate the potential value of the *Privacy Rating*, we conducted a user study. In this early phase of development, we focused on three aspects: the usability of the privacy label, its perceived usefulness and its effect on users' trust in an online service. Due to the COVID-19 pandemic, the data collection took place in individual online sessions. The study was approved by the Ethical Committee of the EEMCS faculty, University of Twente.

7.4.1 Participants

A convenience sample was recruited in three complementary ways: from the university's pool of research participants, from a commercial research participants pool and via social media. Participants from the university's pool received participant credits required by their study programs, participants from the commercial pool received a monetary compensation and participants from social media volunteered to participate without compensation. In our recruitment messages, we called for participants aged 18 or older, with good English proficiency and those in possession of a Google Chrome browser, a webcam and a microphone.

A total of 30 participants took part in the study. Participants had a mean age of 28.6 years (ranging from 19 to 62). Their gender distribution was equal. Participants' educational level varied from medium (high school or vocational education; 53%) to high (bachelor, master, and PhD; 47%). Of the sample, 60% currently followed a study program and 57% had a job. Study programs and occupations were quite diverse. Three participants had a background in cybersecurity or online privacy. All participants lived in Europe, most of them coming from Germany or the Netherlands. A large majority of the participants had ample experience with online tools such as e-mail, search engines, instant messaging, social media and teleconferencing (all 93% or higher) and with online transactions such as online banking, online streaming, and online shopping (all 87% or higher).

7.4.2 Research materials

To evaluate the *Privacy Rating* in a realistic setting, we built an online web shop (see Figure 7.4), using a real, SSL-protected domain (www.sohogiftshop.eu). The web shop used a pre-built, highly-rated WordPress theme. Products (including photos, descriptions and prices) were selected across a broad range. To avoid unintentional visitors, the web shop was password protected. Participants received the password at the beginning of their session.

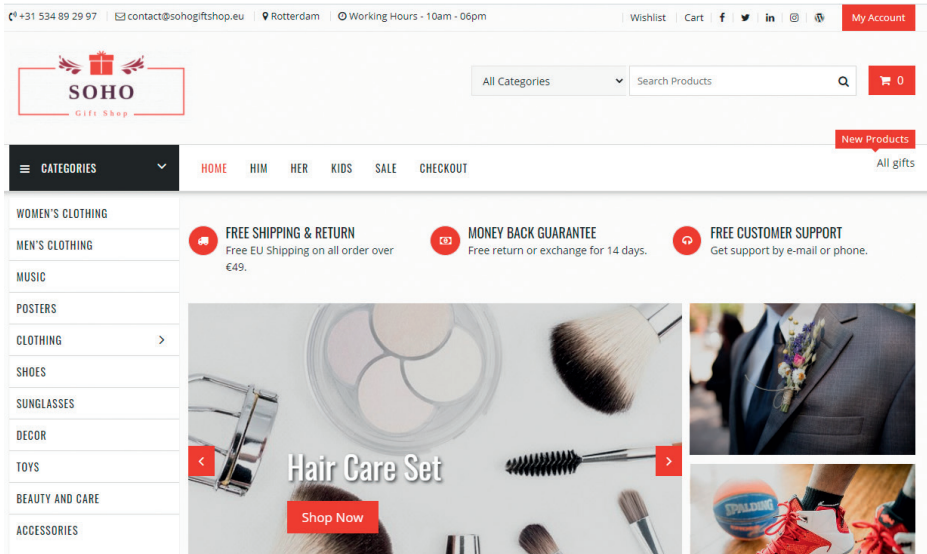


Figure 7.4 Screenshot of the web shop for the user study

The *Privacy Rating* was included in the web shop as a pop-up appearing when opening the homepage. Before interacting with the site, users had to click away the pop-up. The *Privacy Rating* was also included at the bottom of the homepage and a small version was added to the footer of every page (see Figure 7.5).

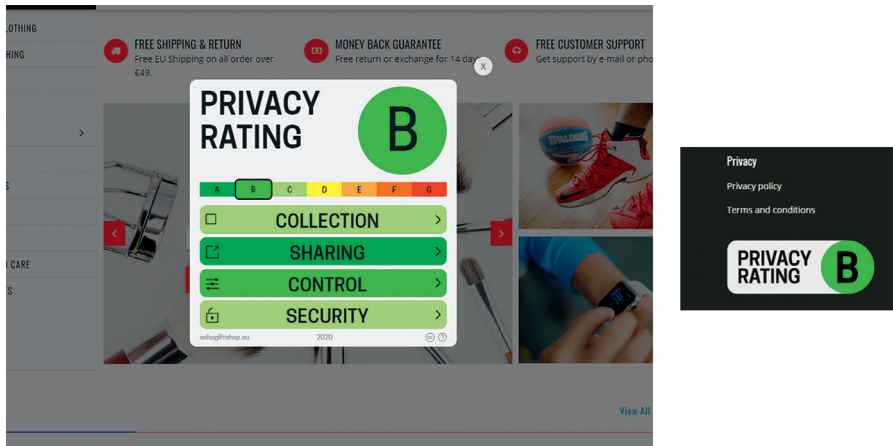


Figure 7.5 The *Privacy Rating* on the web shop, as a pop-up (left) and as small label (right)

To investigate the effects on participants' trust in an online service, two versions of the *Privacy Rating* were used: Half of the participants were exposed to the web shop with a moderately positive privacy rating (grade B, predominantly green) and the other half were exposed to the web shop with a moderately negative rating (grade F, predominantly red).

7.4.3 Procedure

The research sessions consisted of two parts. Participants began with an online questionnaire in Qualtrics covering their background characteristics and the consent information. Background questions focused on age, gender, country of residence, education, profession, use of online services and expertise in online privacy and cybersecurity. After having filled out all questions, participants received a link to a live session with one of the interviewers.

In the live sessions, we used Lookback for real-time screen monitoring and interviewing. Participants were asked to install this software on their computers. In all sessions, two researchers were involved: one moderated the session and interviewed the participant, the other observed.

The session started with a scenario-based task:

“You are looking for a gift for the birthday of your friend. You find some interesting gifts in an online gift shop you have never used before: the SOHO Gift Shop. To place an order, you must provide your first and last name, date of birth, gender, age, shipping/billing address and credit card details. Try to determine whether or not you would trust this website with your information.”

To avoid reactivity, we did not ask the participants to think aloud. However, their interaction with the *Privacy Rating* and the website was recorded and used in the analysis.

The task execution was followed by a semi-structured interview. The questions covered three topics:

- *Trust in the website*: Decision to make a purchase, impression of the website, first impression of the *Privacy Rating*, effects of the *Privacy Rating* on trust.
- *Usability of the Privacy Rating*: Name, overall rating, scale, main categories, detailed information about the categories and visual design.

- *Usefulness of the Privacy Rating*: Transparency (did it increase an understanding of data handling practices?), behavioral intentions (would it affect decisions to trust online services?), desirability (would the participant like to see it as an established standard?).

The sessions were video recorded. Sessions lasted on average 24.4 minutes (SD = 8.3). At the end of the sessions, participants were thanked, debriefed and given instructions for removing the Lookback extension from their browsers.

7.4.4 Analysis

All 30 interviews were transcribed verbatim and any personal information that can be associated with participants was removed. The interview data were analyzed qualitatively in ATLAS.ti. Codes were based on the interview questions and emerged bottom-up based on participants' answers. Two independent researchers coded a random selection of 10% of the transcripts and discussed the discrepancies in their codings. Based on the discussion, the coding scheme was refined. After that, the two researchers coded another sample of the transcripts. They reached sufficient inter-coder agreement, in general (Cohen's kappa = .85) and for the three main research topics: usability (.87), perceived usefulness (1.0) and trust (.78). Using this coding scheme, the remaining transcripts were then coded by the first author.

To investigate the effects of the *Privacy Rating* on participants' trust in the online service, the interviews were complemented with behavioral data: the amount of time participants spent looking at the *Privacy Rating* pop-up and their decision about placing an order in the web shop. For these behavioral data, we compared the results of the two experimental groups (positive versus negative privacy label).

7.5 Results of the user study

7.5.1 Usability

Name. Most participants (80%) found the name *Privacy Rating* clear and understandable and formulated correct expectations of its purpose (“*It’s really clear that this is about how safe a website is in terms of privacy*”). Others stated that they would not know immediately what the name “*is trying to communicate*”. To come to a full understanding, they would have to see more. Interpreting the name in combination with the other elements helped them to “*understand what they mean, what they tell you.*”

Overall Rating. Participants were generally positive (87%) about the clarity of the overall rating (“*It’s understandable enough to make me not want to share my information*”). For most participants, the color was important (“*If there would be no color, it could be like, what does B mean? But green is always good and red is bad*”). Some participants related the overall rating to other familiar grading or rating systems (“*the labels for energy consumption,*” “*the American paper grading system,*” or “*the alphabet; where the alphabet starts, the better it is*”). Participants with difficulties understanding the overall rating stated that the meaning became clearer when they also looked at other elements (e.g., the colored scale).

Scale. Most participants (80%) found the scale clear (“*A would mean that this is the best rating of privacy that you could have as a website and G would be the worst*”). Some called the scale “*intuitive*” and “*nothing to misunderstand.*” The use of colors makes it easy to interpret (“*A is green. So like a traffic light, green is good. Green, you go, you’re safe to go. Yellow as well, you can go....And then red is no, you don’t go. Not very good*”). Some participants stated that a scale without colors would be harder to understand. Others said they need a point of reference to interpret the scale. Interestingly, two participants expected that the scale would be interactive with clickable letters.

Main Categories. The four main categories (Collection, Sharing, Control and Security) were clear to most of the participants, although some participants argued that the terms alone do not suffice and are only understandable when looking at the details corresponding with the categories. The categories Collection and Sharing were easiest to understand (93% and 87%, respectively). Control and Security were somewhat less clear to the participants. Regarding Control (77%), several participants found the term “*a bit vague*” and “*difficult to understand.*” Some thought it refers to the control service providers have (“*maybe what the website can do remotely to your computer*”) and did not see that it is meant to refer to the control users have regarding their personal data. Regarding Security (70%), participants found the term “*a bit ambiguous*” or “*too general.*” Some thought it only involves financial transactions (“*Should I give my Visa number or should I use PayPal*”).

Detailed Information. Most participants (63%) found the more detailed information underlying the four categories clear. Although participants appreciated the conciseness of the descriptions, some suggested adding more information, as it “*is very much open to interpretation depending on how much knowledge the individual has.*” Some participants found the wording too technical and would appreciate explanations in “*more human [layman’s] terms.*”

Table 7.4 summarizes the specific problems participants mentioned with the detailed information.

Table 7.4 Problems identified in the detailed information

Collection	Sharing	Control	Security
What does 'functionality' mean? (n=4)	What does 'legally required disclosure to local law enforcement' mean? (n=5)	What does 'opt-out' mean? (n=7)	What does 'basic security' mean? (n=2)
How long is stored for a limited time? (n=3)	Is not sharing data realistic? (n=1)	What does 'amended' mean? (n=2)	What does 'industry-standard security' mean? (n=2)
Which personal data does it collect? (n=2)			'No anonymization' is a vague term (n=1)
What does data stored indefinitely mean? (n=1)			What does 'legally accountable' mean? (n=1)

Visual Design Elements. Three participants found the separate colors used for the four categories confusing. One participant found the green color difficult to see against its background. Another participant understood this color scheme differently stating that “*sharing and using data are red. So I’m assuming that means that they don’t share my data,*” whereas the color red actually means the opposite. In addition, three participants were confused that the categories expanded both automatically (hover over function) and manually (clickable).

Several participants (33%) did not realize that the indicators were ratings of the single statements (“*The color of the overall rating and the color of the subcategories are the same. I did not notice that those are ratings*”). Another participant thought the colored dots were “*just simple bullet points that don’t have any meaning.*” Especially the green colored dots were hard to recognize, “*because the background is all green and the bullet points are all green.*”

The icons used to support the meaning of the four categories, were correctly understood and appreciated by 44% of the participants. The other participants had difficulties with one or more of the icons. One participant questioned whether or not the icons are really necessary. The interactivity of the icons proved especially confusing for some participants (“*The fact that they move...I can get a little distracted and it makes it look a little less trustworthy to me and not necessarily helping me better understand what it is about*”). Another participant assumed that the icons would be clickable and have a personalization function integrated.

In all, the usability evaluation yielded a positive overall impression as well as several suggestions to further optimize the *Privacy Rating* (see Figure 7.6 for an overview). Some of the detailed problems mentioned with specific elements are actually solved when participants consider the complete visualization. However, the results revealed the need for more attention to the wording of categories and detailed information, with an important balance between clarity and conciseness. In addition, the participants mentioned several ambiguities in the visual design that deserve attention.

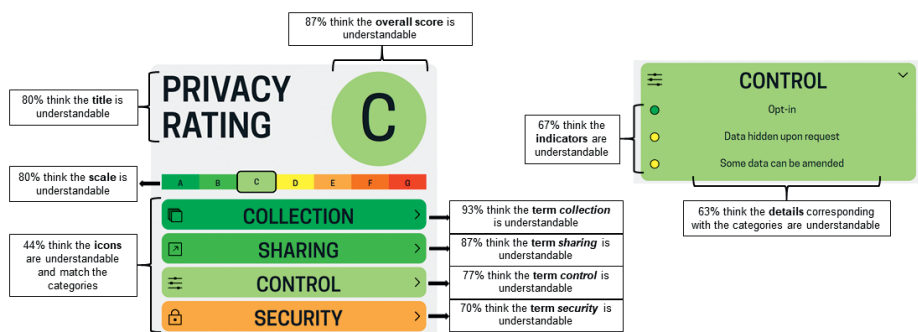


Figure 7.6 Usability of the various elements of the Privacy Rating

7.5.2 Perceived usefulness

Overall, participants were very positive about the *Privacy Rating*; one of them calling it “*the most useful tool I’ve seen.*” The majority of the participants (90%) considered the label to be an effective tool for visualizing how online services handle users’ personal data (“*I think it’s pretty good for a normal homepage because usually it’s not so easy to find this information and I don’t usually read all of it unless it’s a new company*”). The similarities with the existing EU energy label appeared to enhance the label’s usefulness (“*It reminds me a bit of when you buy a fridge and you get the label in terms of the efficiency levels*”). Some participants explicitly appreciated that the label was the first thing they saw when opening a website (“*It gives a pretty clear overview. And it’s also nice if I click on the website and it’s right there*”).

Three participants were somewhat more critical, arguing that the information provided by the label only “*gives an impression but not a full clear explanation of how this website is handling my data.*” Their objections involved the conciseness and clarity of the information, discussed above.

Most participants (83%) felt that the *Privacy Rating* would influence their decisions on trusting and using websites or other online services. They would appreciate such a label, especially when sharing sensitive data such as credit card details with an online service. The label would help them to judge unknown websites or compare services offering the same product. It makes evaluating online services less time-consuming and limits the role of subjectivity in their judgments. Interestingly, some participants argued that a negative rating would influence them more than a positive rating.

All 30 participants would like the *Privacy Rating* to become an established standard under the responsibility of an independent organization as it enhances people's awareness of online privacy and the risks of data sharing, educates users, satisfies the needs of users who care about their personal data and decreases fraud vulnerability:

"I would be happy to see something like that on a website, generally. It would help educate people as to the good and the bad out of the internet, and shopping online and banking online. I surprised myself ... how many online systems I actually use. I worked in IT, but I'd like to think of myself as being able to disconnect from it. But clearly not. Everything I do is connected to technology in some way."

7.5.3 Effects on trust

A first step in our analysis of the effects the *Privacy Rating* had on participants' trust in the web shop involved the attention participants paid to the pop-up. On average, participants spent 33.4 seconds (SD = 18.5) looking at the label (with a range between 6 and 78 seconds). There were no significant differences between the groups who had been exposed to a positive or negative label. In the interview afterwards, almost all participants (93%) indicated to having recognized the label; only two participants were not sure as to whether or not they had seen it.

The second step is to determine whether or not the label affected participants' online ordering decisions. A chi-squared test showed that this was the case ($\chi^2(1, N = 30) = 5.0, p < .05$). In the group of participants exposed to the negative privacy rating, only 40% of the participants would place an order in the web shop, compared to 80% in the group of participants exposed to the positive rating (see Figure 7.7).

Many participants indicated that a negative rating would influence them more than a positive one. Indeed, participants who saw a negative rating placed were less likely to place an order compared to those who were shown a positive rating. Furthermore, while a good privacy rating increased trust in the website for 66%

of the participants, a bad rating decreased trust for 91% of the participants. This indicates that, in our sample, a negative rating had a greater influence on trust than a positive rating did.

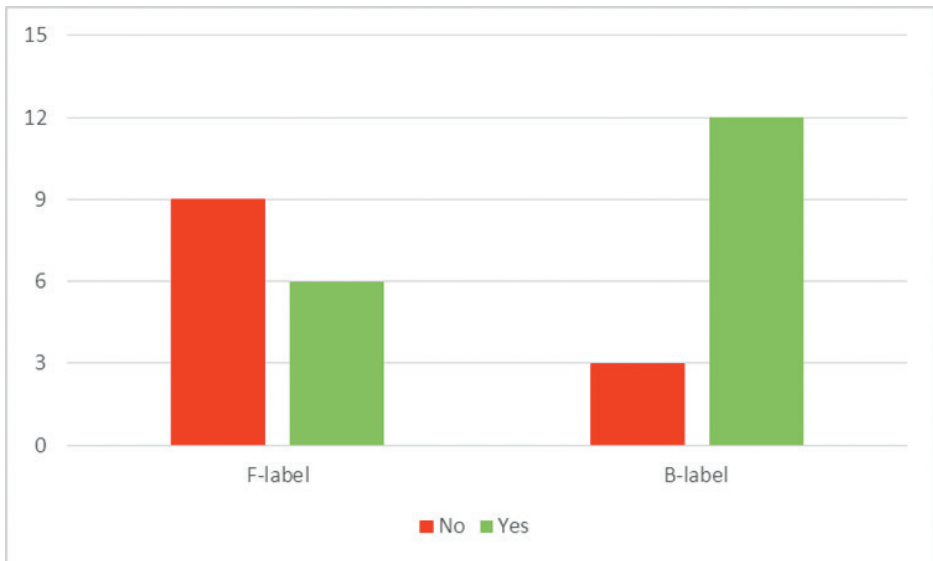


Figure 7.7 Effects of the Privacy Rating on participants' decision whether or not to place an order

From the interviews, two possible factors could be identified that might limit the effectiveness of the *Privacy Rating*. The first is that the pop-up format is not always appreciated. Some of the participants saw it as annoying and disturbing (“I don’t like websites where you have a pop-up straight away....When I go to a landing page of a website, I want to have a look at the actual website and not deal with pop-ups”). The second is that the label is not yet officially established and unfamiliar. This provoked suspicion among some participants:

“I think it’s a bit weird for a website to have that because on other websites that are trustworthy, I don’t see it there....This was a bit unexpected, but unexpected in a negative sense...it could be that they do this in order to make their website look trustworthy while they’re not.”

7.6 Discussion

Online privacy is an increasingly important issue. Rapid technological developments in ICT and artificial intelligence have accelerated the impact of computers and mobile phones in our lives as well as the possibilities for online service providers to invade in our privacy. Interfaces have become deceptively simple and user-friendly, whereas the processes going on in the background are increasingly complex and opaque. Researchers have spent a lot of time and energy unraveling people's privacy-related attitudes and behaviors and exploring the privacy paradox, but so far research-based attempts to empower users with the means to assume responsibility for their online privacy have been limited and unsuccessful.

In this chapter, we described the design and evaluation of a new privacy visualization called *Privacy Rating*. The label acknowledges the functional complexity involved in communicating privacy aspects and supports both less engaged users and privacy-aware users. If widely implemented, it may contribute to privacy awareness among users in general, as it sheds light on the privacy aspects of online services, transforming them from a hidden feature into a conspicuous and comparable characteristic. For less engaged users who may worry about privacy but are unwilling to invest time and effort into evaluating all privacy characteristics, the overall privacy rating provides a visualized shortcut to support their decision-making process as it pertains to the downloading or use of online services. For more engaged users who want to know more about privacy but may be hesitant to examine the entire privacy policy, the *Privacy Rating* offers pre-structured detailed information in two layers. With these contributions, the *Privacy Rating* may play a positive role in balancing the unfavorable equilibrium between users and online service providers, in which privacy considerations currently play an inferior role.

The user research underlined that the *Privacy Rating* can be a promising tool to help users safeguard their online privacy and thus limit the privacy paradox (Barnes, 2006). Of the three underlying mechanisms of the privacy paradox—a more or less rational weighing of costs and benefits, an incomplete and biased weighing of costs and benefits—and a neglect of privacy considerations (Barth & De Jong, 2017)—it helps reduce the influence of the latter two. The label urges users to consider privacy aspects in their decisions and reduces biases they might have when judging privacy risks. As a result, the weighing of costs and benefits will be more systematic and more rational than may currently be the case. That does not mean that the privacy paradox is solved. It is still imaginable that users decide in favor of a certain online service, despite the privacy risks they are aware

of. But the discrepancy between attitude and behavior may not be at the core of the problem. People have to make trade-offs between desires and preferences all the time. The core of the problem is the fact that their decisions are often uninformed. Tackling this deficit is the main purpose of the *Privacy Rating* tool. The results of the user research suggest that the design used can be considered a step in the right direction. With regard to usability, the *Privacy Rating* did quite well, although participants also uncovered several problems that need to be addressed in future iterations of the label. The problems found mainly concerned the formulation of privacy risks and aspects and details in the visual design. The perceived usefulness was judged very favorably by our users and the label appeared to significantly affect our participants' decisions on whether or not to use a particular web shop. User feedback will play a significant role in our future efforts to further optimize the *Privacy Rating*.

In addition, the results of our user research can be used to inform other privacy visualization projects. Two insights stood out. The first is that it is beneficial if a privacy visualization explicitly connects to existing interpretation frames of users. In all parts of the user research we heard positive remarks about the resemblance of the *Privacy Rating* with the well-established and familiar energy label, which made the label easy to understand and may also have contributed to the persuasiveness and perceived urgency of the rating. The second is that the development of the label is only half of the story. Several participants in the user research doubted the independence and authoritative nature of the label, letting on that it would make a big difference to them if the label is issued by a trustworthy source.

Finally, our findings drew attention to two trade-offs in designing a privacy visualization. The first involves finding a balance between conciseness/simplicity and informativeness. The feedback from some of our participants suggested that they found even the second layer in the information about privacy insufficient. Having said this, it is by no means certain that adding information will make the label better. Our findings lend support to previous work stating that grouping and segmenting information across multiple layers has a positive effect on the understandability of complex information (L. Edwards & Abel, 2014) and that color schemes can increase granularity and provide shortcuts for quickly assessing risks (S. De Jong & Spagnuolo, 2020; Efroni et al., 2019). Also in line with previous work, we found that privacy and security icons have poor understandability (S. De Jong & Spagnuolo, 2020; Rossi & Palmirani, 2017). The second trade-off is between annoying intrusiveness and sheer invisibility. Some of our participants complained about the use of a pop-up, but it is questionable whether or not a less intrusive exposure would glean the necessary attention and

provide similar effects. Earlier research showed that the timing of users' exposure to privacy notices is very important (Balebako et al. 2015). The development of any viable privacy visualization must include the effective placement of such. It is quite possible that the methods of exposure may become less important once the label becomes an established standard (Esayas et al., 2016).

7.6.1 Practitioner's takeaway

- This study describes the design and evaluation of a privacy visualization (*Privacy Rating*) aimed at empowering users to protect their online privacy.
- Functional complexity is a major design challenge: Empowering users implies making them aware of privacy risks, and giving them shortcuts as well as access to more detailed information in a clear, concise, and intuitive design.
- User research shows that the *Privacy Rating* fulfills the needs of users: Usable and useful, it significantly influences users' trust in online services.

7.6.2 Limitations and future work

To our knowledge, this is the first initiative to develop a privacy visualization covering a systematic selection of relevant privacy attributes available in academic literature, law and practice. It is also one of the few initiatives to explicitly incorporate user feedback in the process. Still, it is important to keep the following limitations in mind when interpreting the results.

First, the *Privacy Rating* is still in development. In our user study, we tested a prototype of the privacy label, which reflected our knowledge after various studies into user perspectives on online privacy (Barth, De Jong, et al., 2020; Barth, Ionita, et al., 2020; Barth, Ngo, et al., 2020), after a thorough analysis of relevant privacy aspects and earlier privacy visualizations (Barth, Ionita, et al., 2020) and an iterative design process including expert and user input. The user study reported in this article provided us with more food for thought, which we will use to further optimize the *Privacy Rating*. Concretely, we will look into using simpler language and including links to further information. In addition, we foresee three extra developments in the period ahead. We will try to further explore the implementation of the label, which involves gaining support from online service providers, platforms and/or legislation. Any advancements may have consequences for users' perceptions of the *Privacy Rating*. Moreover, we will try to make the input for the *Privacy Rating* more objective and trustworthy. The score online services get is now based on service providers' self-reports in the questionnaire. That is not necessarily a bad option, as service providers

can be held responsible for any discrepancy between their privacy policies and their answers in the questionnaire. But ideally, the privacy ratings would be obtained directly from the privacy policies, either by natural language processing or by the intervention of an independent authority. Future developments in this respect may also have a positive impact on users' perceptions. We will try to set up communication about the *Privacy Rating* itself. In the current user study, participants saw nothing but the visualization. We are planning to develop a series of short persuasive messages explaining the system, its background and its necessity.

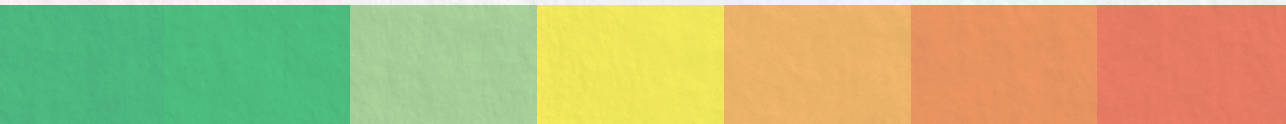
Second, the user research described in this article was, in line with the state of development of the *Privacy Rating*, limited to specific aspects of the label after artificial exposure. In the usability test, we focused predominantly on the perceived understandability of the various elements of the *Privacy Rating*. It would be interesting in follow-up research to focus more on participants' interpretations and actual use of the label as a whole. The research into the effects of the *Privacy Rating* was limited to the explicit question as to whether or not the participants would trust the web shop enough to do business with it. Follow-up research in a more natural setting would less exclusively and explicitly focus on the trust question and also include how the privacy label might, for instance, affect the image or reputation of the online service provider. The question as to whether or not a positive privacy rating can be good for business would be very relevant, as it could convince online service providers to embrace transparency regarding privacy and include the *Privacy Rating* in their communication.

Third, experimental research is needed to further investigate the two trade-offs we mentioned: between conciseness/simplicity and informativeness (which balance is most effective for which user groups?) and between annoying intrusiveness and sheer invisibility (how can we make a privacy label optimally visible without annoying users?). And finally, it would be interesting to extend research in laboratory settings with real-life research into the users' appreciation of and behavior toward the *Privacy Rating*.

7.6.3 Conclusions

We propose *Privacy Rating* which addresses the inherent functional complexity of privacy communication by visually synthesizing information across multiple layers of increasing detail. It thereby increases awareness, provides less engaged users with shortcuts and supports privacy aware users in making informed decisions. Usability testing showed the label was perceived as useful and usable. It had a significant effect on trust in the online service. All participants indicated they would appreciate such a label becoming an established standard.

More generally, we learned that privacy visualizations should use familiar design elements and ideally be supported by a trustworthy organization. A good privacy visualization should be both concise and informative. Providing visual shortcuts by means of layering information and using well-known color schemes help; icons do not. Furthermore, the visualization should be placed so as to be noticed but not annoy. A privacy visualization which satisfies these requirements can empower users by significantly improving privacy awareness and helping achieve truly informed consent.



8

General discussion

“Although rhetoric often names the technologies themselves as sources of concern, e.g. “big data,” or “biometrics,” the sources of privacy threats are...technologies embedded in particular environments shaped by social, economic, and political factors and practices and put to specific purposes. Most salient to individuals are practices of familiar actors with which they are directly acquainted, such as Facebook, Google...More informed critics point to information brokers, backend information services, ad networks...and biometric identification systems...which relentlessly monitor and shape lives in ways neither perceptible nor remotely comprehensible to the public of ordinary citizens” (Nissenbaum, 2018, p. 832).

8.1 Main findings

Using the internet safely is by no means a simple or obvious matter. Should third-party cookies be accepted in their entirety, or only those necessary for website functionality? Are random or personalized ads better? Is a secure password sufficient or should two-factor authentication be enabled? Which browser provides sufficient privacy? Is the website encrypted? Which e-mail provider is private? The list of unknowns and uncertainties is practically endless when using online services. The information age and the shift to online environments for a wide range of activities means that users are confronted with important and increasingly difficult data exchange issues. Given the ‘choice’, most users opt to preserve their privacy rather than allow those settings considered intrusive. This is the crux of the matter: *Do users understand privacy settings and policies? Do they possess the required skills, mental capabilities, necessary time and willingness to take control of their privacy online? Furthermore, are they aware of the consequences of their day-to-day actions online?* According to our research, the answer to all of these questions is a resounding ‘No’.

The possible explanations for this issue are complex, and the solutions cause debate. Privacy preferences are shaped by a sociotechnical system (Nissenbaum, 2010), but individual characteristics, attitudes toward online privacy and trust in the actors involved in the data exchange also play a role. Furthermore, the centralized and take-it-or-leave-it nature of online services built upon invisible data highways creates asymmetric power relations between the data subject and data holder. Users end up disclosing data despite privacy concerns and the need to preserve their privacy zones in the online environment (Arkko, 2020; Berendt, Günther, & Spiekermann, 2005; Bräunlich et al., 2020; Kelbert et al., 2012; Müller et al., 2012; Ochs & Löw, 2012). This contraction between stated preferences toward online privacy and actual disclosure behavior is heavily discussed in the academic literature. While several scholars explain this ‘privacy paradox’ from a rational trade-off perspective whereby perceived benefits outweigh perceived risks (Keith et al., 2013; Y. Li, 2012), other scholars seek the explanation for this indifferent behavior in cognitive limitations and biases, heuristic thinking, or insufficient interest in the topic of privacy (Acquisti, 2004; Deuker, 2010; Keith et al., 2013; Shklovski et al., 2014). To date, the controversy continues, and there is no universally accepted theory, no consensus about the mental processes that guide decision-making or a solution to close the gap between privacy preferences, needs and actual information disclosure.

The overall aim of this dissertation was to develop a research-based approach toward empowering online users by ensuring that they are comprehensively

informed about the data handling practices of the online services they utilize. Ultimately, educating users about data handling practices will enable them to better protect their privacy by tackling the crux of the online privacy problem: helping users understand privacy settings and policies in a way that requires a minimum of cognitive involvement, time and digital skills, while making users aware of the risks pertaining to information disclosure so that a willingness to take control of privacy online can be fostered. To achieve this, this dissertation had two research goals. The first research goal centers around knowledge acquisition and gaining insights into the online privacy behaviors of users to better understand what factors drive information disclosure. To this end, a literature review and three empirical studies were conducted. The second research goal centers around a design approach aimed at ascertaining a viable solution to visually communicating the most relevant aspects of a privacy policy to users. For that purpose, another literature review was conducted that served as input for the design of a user-centered privacy visualization.

First, to identify theoretical approaches that explain the disparity between stated interest in online privacy and actual disclosure behavior, a systematic review of the existing literature on the privacy paradox was conducted (**Chapter 2**; research goal 1a). Based on a sample of 32 full papers covering 35 theories, an overarching theoretical framework was developed, addressing the discrepancy between stated privacy concerns and protective behavior through different theoretical lenses. The theories were grouped into three distinct types of decision-making as it pertains to information disclosure online: (1) a *rational* risk-benefit calculation, whereby perceived benefits outweigh privacy threats, (2) a *biased* risk-benefit calculation, affected by nonrational factors or bounded rationality and (3) a decision-making process with *no* or only *negligible* risk consideration. Biased risk-benefit calculations such as heuristics thinking, immediate gratifications, or habits (Acquisti, 2004; Debatin et al., 2009; Deuker, 2010; Gambino et al., 2016) and a superficial or absent risk assessment because privacy valuation fails or information asymmetry prevents it (Flender & Müller, 2012; Oetzel & Gonia, 2011) are the most common explanations found in the analyzed theories. This finding indicates that the privacy paradox is more than “the perfectly rational pictures we are used to build” (Dinev, 2014, p. 100). The (near) impossibility of assessing risks to personal data and the intangibility of privacy breaches often results in a state of uncertainty (Acquisti, 2009; Acquisti et al., 2017). Having incomplete information about factors that might play a role in an online environment makes it very unlikely that the data disclosure situation is assessed rationally (Dinev, 2014; Harsanyi, 1967). Rather, it is more than likely that bounded rationality, biases and heuristic thinking guide the decision-making

process. Even if users have privacy concerns on a general level, factors such as low transparency, unfriendly user design, and unfair privacy policies constrain protective online behavior.

To go beyond theoretical explanations, an experimental study aimed to investigate the privacy paradox ‘in the wild’ was conducted (**Chapter 3**; research goal 1b). The strength of this study is that actual behavior was measured and that the context was taken into account. Research often cites a lack of technical knowledge, limited privacy awareness, and financial considerations as the causes for the privacy paradox. Therefore, the study aimed to control for technical expertise and financial considerations by studying a tech-savvy user group and providing monetary compensation for the purchase of a mobile app. Furthermore, it is said that users relativize their privacy attitudes in the heat of the moment (Müller et al., 2012). Therefore, participants were asked before and after the actual installation process which factors played a role in the decision-making process as it pertains to selecting, downloading and installing an app. Although participants indicated they take permissions into account when downloading an app and that trust in the app plays a significant role, the results showed that privacy and security considerations were outweighed by price, ratings and design. Despite expressing concerns about unauthorized access by third parties, most participants were not willing to spend money on a comparable app that did not ask for unnecessary permissions. The results indicate that technical knowledge does not shield individuals from biases in decision-making—as other attributes such as design and/or ratings play just as important a role in the decision-making process when downloading a mobile app. Furthermore, the fact that participants paid considerably less attention to permission requests than previously claimed prompts the question regarding whether or not even subjects with a technical background understand enough about permissions and their potential ramifications. The latest version of the Android permission system only requests permission at runtime in an effort to minimize information overload. However, it is highly questionable whether this objective has been fulfilled. It could be argued that if a user has already selected and committed to downloading an app, it is likely that little attention will be given to permissions requested at runtime. This results in the acceptance of potential privacy intrusions despite privacy concerns on a general level.

To delve further into the effect of technical knowledge and privacy awareness, an interview study with experts working in the privacy and cybersecurity sectors was conducted (**Chapter 4**; research goal 1b). An argument often put forward to explain the contradiction between stated privacy preferences and actual information disclosure is the lack of necessary technical knowledge. In

response to this argument, this interview study aimed to examine how privacy and cybersecurity experts deal with their online privacy. Analysis showed that participants could be evenly divided into three groups according to their attitudes toward privacy and their reported behavior regarding mobile phone usage: (1) experts that (highly) value their privacy and are concerned about the loss of their personal data, (2) experts that value their privacy but that are not overly concerned, and (3) experts that do not pay much attention to the protection of their personal data. Interestingly, these three groups correspond to the categories of Westin's Privacy Index Segmentation (Westin, 1967): privacy fundamentalists, privacy pragmatists and unconcerned users. Although experts identified as privacy fundamentalists seemed to highly value their personal data, their online service usage did not overly reflect this, as they even used risky apps that ask for permissions that are not directly related functionality. They justified the contradiction between stated attitudes and behavior by time constraints, group pressure, simple convenience or the desire to use an app. To diminish cognitive dissonance, concerned experts try to evaluate the app's permissions according to the functionality of the app, although they are sometimes hard to interpret. This is in contrast to experts belonging to the group of privacy pragmatists. These experts are well aware that personal data are not always treated confidentially, and although they indicated a sense of unease with the situation, they were still willing to use such online services, failing to adequately protect themselves against privacy intrusion. Whereas unconcerned experts claimed to know about the risks to personal data, they had no objection against information disclosure online in general. Based on these results, the main findings from the interviews are threefold: (1) technical knowledge does not automatically lead to more cautious privacy-related behavior, (2) the justifications of risky online behavior do not differ between experts and general users, and (3) even if experts review different cues than their general user counterparts, the resulting online behavior does not significantly differ. In conclusion, expert users are as vulnerable to heuristic thinking and cognitive biases as general users.

To gain insight into users' privacy perceptions and preferences, a study with general users was conducted (**Chapter 5**; research goal 1c). More specifically, using the Q-sort method, groups of users were segmented based on their views toward privacy preferences. Inferred from the theory of contextual integrity (Nissenbaum, 2004, 2011), privacy perceptions were assessed under consideration of the type of app (health vs news app) and the three contextual factors: the type of personal information collected (*what*), how the information is processed (*how*) and the party involved in the information transaction (*who*). Interestingly, the type of app seems to play only a contingent role in users' privacy perception. The

primary focus when forming privacy perceptions was on the type of data that was collected and what was done with the data. Furthermore, actors involved in the data exchange process received limited attention. These findings provide partial support for the theory of contextual integrity as the type of app (context in which the data exchange takes place: *why*) and the actors involved in the data exchange process (*who*) did not appear to be important when forming privacy preferences. However, the type of collected data (*what*) and the way information is processed (*how*) play a major role in users' privacy perception. With regard to these two contextual factors, users differ in their views on acceptable practices, whereas privacy infringements were perceived largely the same across all groups. In other words, users largely agreed on what is perceived as a strong violation of their online privacy in terms of *what* information is collected and *how* this information is processed.

The second research goal concentrates on the question of how to visually communicate the most relevant aspects of a privacy policy to users. First, to address this question, a further literature review was conducted. The literature study aimed to systematize knowledge from academia, industry, and government to identify generally applicable privacy attributes of online services (**Chapter 6**; research goal 2a). For this purpose, existing privacy visualizations and Privacy by Design (PbD) guidelines published in the last 19 years were reviewed and analyzed. Eventually, this systematic review resulted in a unified list of 15 privacy attributes: accountability, anonymization, collection, control, correctness, disclosure, functionality, pseudonymization, purpose, retention, right to be forgotten, sale, security, sharing, and transparency. With the help of privacy experts, this list was validated and refined, and definitions were agreed upon. The results revealed significant differences in attributes covered by privacy visualizations and PbD guidelines. To take into account the perception of users regarding these attributes, 385 users and 100 privacy experts were asked to rank the 15 attributes according to perceived importance. The results showed that the collection, sharing and sale of data were perceived as most important by both users and experts. However, all 15 attributes were ranked on average high, a fact that does not allow the prioritization of one attribute above another. Interestingly, the sale of data was seldom present in PbD guidelines and did not appear in privacy visualization, despite having high importance to users. Accountability and anonymization were also rarely included in visualizations, and the right to be forgotten was often neglected by PbD guidelines, even though these were rated highly by users. This leads to the conclusion that existing privacy visualizations are not yet user-oriented. Despite these shortcomings, initiatives for developing PbD guidelines and privacy visualization increased after 2009

and 2012, respectively. PbD guidelines were predominately approached from a regulatory point of view or were introduced as industry standards, indicating that privacy by design had entered practice. Moreover, communicating privacy to users and raising awareness of this topic by means of privacy visualization gathered steam. However, despite increasing initiatives for implementing privacy into design and empowering users as mandated by the GDPR, no official standard or generally accepted privacy label has yet been established.

To address this shortcoming, the last study proposes a user-validated privacy visualization (**Chapter 7**; research goal 2b). The unified list of privacy attributes found in the literature from academia, industry and government was taken as a starting point for the privacy visualization. From these 15 attributes, 12 privacy attributes were eventually included in the rating system. Anonymization and pseudonymization were merged, functionality was excluded because it closely resembles control, and transparency was excluded because a visualization represents transparency in itself. In the next step, the 12 attributes were divided into four main clusters: *collection, sharing, control, and security*. Each attribute was made measurable on a three-point scale (good-neutral-bad) corresponding to a penalty scheme ranging from 0 points for good scores, 1 point for neutral scores and 2 points for bad scores. Eventually, the total number of penalty points (best score = 0, worst score = 24) was categorized into seven classes, from class A (lowest privacy risk) to G (highest privacy risks). These classes correspond to a color scheme from green (class A) to red (class G). The aim of the privacy visualization, called the *Privacy Rating*, is threefold. First, the visualization aims at enhancing awareness of the online privacy topic in general. Second, using an overall privacy class in letters in combination with colors helps users who are less engaged with their online privacy to obtain a quick and easy-to-access overview about an online service's privacy risks. Third, users who are more interested in their online privacy are helped with two further levels that provide more specific information about the four categories (collection, sharing, control, and security). Eventually, we showed that it is possible to design a privacy visualization that presents a complete overview of the most important aspects of a privacy policy in a user-centered fashion without risking information overload while satisfying the information needs of different user types. The results from the user study confirmed this conclusion, as the *Privacy Rating* was well understood by the users in our sample and had a significant effect on the perception of trust. Overall, users emphasized the simplicity, clarity and attractiveness of the design. Furthermore, all participants stressed the importance of such a rating and expressed a desire to see this kind of label as an established standard. In conclusion, these findings show that using visual stimuli

and simplified descriptions of otherwise complex privacy policy information can motivate users to invest more interest in how their personal data are handled. This brings the overall goal of this dissertation full circle: the empowerment of users to make well-informed decisions on what information they are willing to disclose with whom and for which purposes. Simply put, give users control over sharing (Acquisti et al., 2016; Passera, 2012; Sheng et al., 2020).

8.2 Theoretical implications

The information age is characterized by technologies and devices that are increasingly interconnected. They facilitate access to a vast amount of information and allow large-scale data aggregation and the analysis of such information. Activities that originally took place in an offline setting are increasingly shifted to the online environment, raising serious privacy and security issues. For instance, round-the-clock social networking connectivity, online dating, online shopping or online information seeking now plays a part in the everyday lives of many. Consequently, along with information technology, separating the public space from the private space becomes increasingly difficult. As a result, users are regularly forced to make complex decisions regarding what personal information (not) to disclose and to whom. This dissertation makes the following theoretical contributions to the understanding of information privacy in an online environment.

8.2.1 Information disclosure is not a rational risk-benefit calculation

The systematic literature review of theories explaining the privacy paradox resulted in three categories of decision-making that explain the discrepancy between stated privacy preferences and actual disclosure behavior: (1) a rational risk-benefit calculation, (2) a biased risk-benefit calculation, and (3) a situation in which no or little risk assessment takes place (Chapter 2).

Surprisingly, privacy calculus theory (Culnan & Armstrong, 1999)—whereby the intention to disclose information is guided by a rational risk-benefit calculation—is often used to explain perceptions regarding information disclosure. However, the findings of the literature study of Chapter 2 and the biased decision-making observed in both tech-savvy students (Chapter 3) and cybersecurity experts (Chapter 4) provide no supporting evidence for such a rational calculation. In fact, Dinev (2014) and Veltri and Ivchenko (2017) see decision-making in an online environment as far from rational. This is primarily because the likelihood and impact of privacy risks are hard to estimate for users

or users simply do not know what the risks are (Acquisti, 2004; Acquisti et al., 2015; Bräunlich et al., 2020; Flender & Müller, 2012). It is also difficult to identify and compare alternatives (Simon, 1982, 1990). Furthermore, the online environment is characterized by fast technical processes and opaque structures of data highways that leave individuals in a state of uncertainty (Acquisti, 2009; Acquisti et al., 2017), where the user often knows less about data handling processes than the data processor. As such, privacy perception as a rational weighing of benefits and risks (Dinev & Hart, 2006) is limited by information asymmetry (K. Martin, 2013). Even if all of the necessary information were easily available, the complexity of online data handling and its underlying technical infrastructure would often exceed the cognitive abilities of users. When faced with incomplete or overly complex information, individuals often resort to heuristic thinking (Simon, 1990). Unfortunately, the bounded rationality of the data subject seems to result in a situation of information disclosure to the detriment of the user. It is therefore crucial that online privacy be understood in terms of bounded cognitive capabilities and psychological factors, technological developments and characteristics of the technology itself.

8.2.2 Knowledge and privacy awareness do not play a role in information disclosure

In the context of the privacy paradox, it is often assumed that a knowledge and awareness gap leads to indifferent online behavior (Bandara et al., 2017; Liccardi et al., 2014; Volkamer et al., 2015). Arguably, users who are poorly informed about data handling practices and the consequences of information disclosure are more ‘careless’ in sharing information online (Acquisti et al., 2015, 2016; Bräunlich et al., 2020; Shklovski et al., 2014). It is therefore fair to assume that technical expertise or an interest in online privacy would lead to different evaluations of information disclosure and more cautious behavior (Ion et al., 2015). However, the studies conducted among a technically literate sample (Chapters 3 and 4) do not support this hypothesis. Interestingly, both the privacy evaluation and the reported and actual online behavior of technical experts is comparable to that of general users. Their expertise seems to be overridden by situational cues and internal considerations. Although experts indicated that they had privacy concerns and an understanding of the potential consequences of data disclosure, they were found to be vulnerable to heuristic thinking, biases and the temptation of immediate gratification, as were their lay counterparts. This finding is in line with those of De Luca et al. (2016), Debatin et al. (2009), and Kang et al. (2015), all contradicting the common assumption that digital literacy and privacy awareness result in better protection of personal data. However, relativizing the

influence of (technical) knowledge and heightened awareness as they pertain to the privacy threats associated with information disclosure does not simplify the user empowerment issue in general.

8.2.3 Online privacy is multifaceted

Online privacy is subjected to technological change and developments and should therefore not be considered as a fixed concept but as a process that is in flux (Trepte, 2016). Essentially, online privacy centers around the concepts of access and control. To execute privacy protection, users must be in control over who has access to their personal data. However, allowing or denying access to one's data is not a fixed decision but changes with the context, social and situational boundaries and norms. What is deemed to be acceptable to disclose in one situation might be perceived as damaging in another. This fluency of privacy boundaries makes online privacy so hard to define (K. Martin, 2016a; Nissenbaum, 2004; Pavlou, 2011, Solove, 2008). However, the results from the Q-sort study of Chapter 5 showed that many aspects fall under the term 'online privacy' and that users have different views on what is deemed private information. Without a doubt, privacy is shaped through the sociotechnical system (Nissenbaum, 2010), but privacy preferences are also formed by individual attitudes. Interestingly, the Q-sort study revealed that users agreed upon unacceptable practices that are considered violations of information privacy, such as unauthorized access to private photos or messages or the selling and sharing of data without consent. In addition to agreements on generally unacceptable practices, users can be grouped according to practices they are willing to accept when using online services. One might be concerned about personal identifiable information but accept disclosure about hobbies in general, whereas another would withhold that information because of profiling concerns (Müller et al., 2012). Therefore, to understand privacy in its totality, context dependency needs to be taken into account, focusing on 'core characteristics' of online privacy and considering norms and values guiding information exchange (Nissenbaum, 2004; Solove, 2002, 2008). The theory of contextual integrity can be considered a promising approach to reduce confusion and ambiguity about information privacy and ultimately benefits privacy protection, as "respect for context means consumers have a right to expect that companies will collect, use, and disclose personal data in ways that are consistent with the (social) context in which consumers provide the data" (Nissenbaum, 2018, p. 850).

8.2.4 User empowerment is needed

The GDPR (The European Parliament and the Council of European Union, 2016) and the U.S. Federal Trade Commission (FTC; Anthony, 2001) strongly recommend visually communicating online privacy to users. However, no official standardized privacy label has yet been developed. Furthermore, although the GDPR mandates PbD, the findings of Chapter 6 are in line with Bygrave (2017) in that the meaning and implementation specifics of PbD are still unclear. It seems that existing PbD guidelines and privacy visualization disagree in regard to what constitutes ‘good privacy’ (Chapter 6). Significant differences were even observed between sets of PbD principles in terms of both content and granularity. Moreover, when making decisions on information disclosure and whether to use an online service, users are often vulnerable to heuristic thinking and cognitive biases (Chapter 2). To make the problem even worse, technical knowledge and general risk awareness do not prevent biased decision-making (Chapters 3 and 4). At the same time, users indicate that a lack of knowledge and privacy awareness lead them to questionable disclosure behavior. However, if prompted by a privacy awareness tool, users appeared to be willing to deal with their online privacy in a responsible manner (Chapters 5 and 7). These results clearly show that users want to have control over their personal data, but they need support to act accordingly. In line with Baek (2014), Deuker (2010) and Pötzsch et al. (2010), a privacy awareness tool at the situation-specific level that communicates the information handling practices of online services has a realistic chance of making users less vulnerable to disclosure influences, biases and heuristic thinking. Although expert knowledge about information privacy seems to influence decision-making to a lesser extent than expected, enhancing privacy awareness at the situation-specific level and at ‘the heat of the moment’ is a promising approach to span the gap between privacy preferences and online disclosure behavior. Eventually, as requested by the GDPR and shown by the results of this dissertation, visually communicating the privacy standards of online services is needed. Informing users in a concise and easy-to-understand manner about the privacy risks of an online service will empower users to protect their online privacy while balancing the scales between privacy preferences and the potential risks associated with information disclosure.

8.2.5 The main difficulties with online privacy stem from communication problems

The results from the literature review presented in Chapter 2 and the empirical studies conducted within the scope of this thesis (Chapters 3-5) revealed that users of online services can be indifferent in their information disclosure

decisions. They disclose information despite having privacy concerns, and privacy preferences vary from one situation to another. Moreover, there is no single unilateral perception of online privacy. Rather, users can be grouped based on their general privacy perception (e.g., concerned, pragmatic or unconcerned) but also on context- and situation-specific preferences on what data are deemed acceptable to disclose and for what purpose (e.g., anonymous data, tracking, against profiling, or in favor of personalizing). Looking at the historical development of technological innovations, applications and services have become increasingly more user-friendly and intuitive, while the technologies—and the associated processes powering them—have become complex (Lowdermilk, 2013). The general misunderstanding of the Internet as a free service, lengthy privacy policies, annoying cookie notices and the opaque and complex smartphone permission system exacerbate the problem of careless online behavior (Antón et al., 2004; Benton et al., 2013; Y. Chen et al., 2019; Fernback & Papacharissi, 2007; Kelley et al., 2012; Kucuk, 2016). Altogether, individual privacy preferences, fluent privacy boundaries, and the context dependency of privacy combined with poorly implemented privacy enhancement measures sustain the functional complexity of information privacy: Online privacy communication unites multiple goals for multiple stakeholders. First, online services and their practices need increased transparency. Second, all users—regardless of their privacy perceptions and preferences—need to be made aware of important privacy issues. Third, users—the interested as well as the less interested ones—must be supported with an easy-to-understand and fast-to-process solution that facilitates well-informed decision-making as it pertains to their online privacy. Our research leads us to believe that these three main issues must be effectively addressed if universally acceptable privacy standards in online environments are to be achieved.

Complicating matters even further, previous research on communicating complex information showed that exhaustive documents do not work (Beldad et al., 2010). However, online privacy is an overly complex topic. Hence, online privacy is an interaction between many factors influencing each other, prompting questions on how to communicate privacy topics. The design approach presented in Chapters 6 and 7 showed that developing a system for operationalizing and measuring the privacy of online services is possible. Moreover, the look-and-feel combination of known design elements (e.g., color scheme, resembling the EU energy label) with a multilayer approach to satisfy the different information needs of users covers all relevant privacy aspects in a user-friendly way.

8.3 Practical implications

This dissertation provides actionable insights into the perceptions and behavior of both users and experts with regard to online privacy. It also provides a user-tested, production-ready privacy rating and visualization system. Eventually, the goal is to empower users of online services to relieve the tension between privacy preferences and the risks associated with information disclosure and to pave the way for improved online privacy protection. Reaching this goal requires the involvement and collaboration of several key stakeholders: researchers, policymakers and practitioners.

First, it is important that all stakeholders start to understand online privacy as a multifaceted concept. Online privacy is not a specific class of information. It is not a concept that can be simply approached and protected by regulations only. This dissertation shows that not only users and privacy experts disagree on what factors play a role in the concept of online privacy but practitioners, researchers and regulators also remain divided. Moreover, the importance that is given to types of personal data varies considerably among the users of online services and the situations in which technology is used. Therefore, privacy needs a multidisciplinary approach incorporating views from communication, legal, technical, economics, psychology, and political science (Dinev, 2014; Pavlou, 2011). The actors involved in the data exchange, the type of information exchange, what is done with the data and how they are handled are complementary to each other and must be seen as a whole (Nissenbaum, 2018). Moreover, the technology itself that functions as a mediator for information exchange must be taken into consideration. This requires taking a bottom-up approach and defining what a technology does, as well as determining potential disruptions to privacy, e.g., what are potential scenarios for a breach of confidentiality (Solove, 2002). Therefore, and in line with Bräunlich et al. (2020), Buchmann (2012), Nissenbaum (2010) and Solove (2008), the understanding of online privacy should not only be guided by the stance held by users and experts, but also approached from various disciplines. Interdisciplinary collaboration between legal, technical and social science under the consideration of the industry is needed. To paraphrase Solove: “the need to conceptualize privacy is significant, but the discourse about conceptualizing privacy remains deeply dissatisfying” (Solove, 2008, p. 2).

Second, we live in a data economy that entails data gathering to keep that economy running. Therefore, privacy must not only be understood in terms of contextual integrity but also be seen as a business model (Acquisti, 2010; Nissenbaum, 2011). The guiding factor behind this business model should be

transparency and a balanced power relationship between the data subject and data holder. This eventually means that service providers must grant the user incessant control of their personal information. This means shifting some of the responsibility away from the data subject. More specifically, online services should extend the concept of corporate social responsibility to corporate privacy responsibility (Bandara, 2020). This also means that information exchange in an online environment should not only be understood as a purely legal contract expressed in a privacy policy but also characterized by a social contract based on generally expected moral values. Eventually, a fair legal and social contract between the data subject and data holder is the only viable way to achieve a trustworthy environment in an online world that is becoming increasingly interconnected.

Third, to satisfy the principles of notice and choice, educating the user about data handling practices is of utmost importance (K. Martin, 2013). The findings gained from the design-based approach showed that various metrics of information privacy could be identified. Moreover, it is possible to operationalize and objectively measure those privacy attributes to show the privacy level of an online service. This not only informs users about the data handling practices of a specific service but also allows them to compare different services. The rating we propose has the potential to function as a standardized, understandable, and machine-readable summary of a privacy policy. Although we believe communicating data handling practices to be crucial to maintaining a viable provider-user relationship, this aspect of the privacy ‘problem’ seems to be strongly underrepresented in the technical and professional communication discipline. Moreover, to bring the *Privacy Rating* to the market, alliances between various stakeholders from industry, science and regulatory bodies need to be formed, ideally on a European level, to make the label an established standard. Similar to the European energy label, a privacy visualization has to become mandatory before it can gain general acceptance by service providers and users alike. This is an ambitious goal, but this dissertation showed that it is possible to design a user-friendly privacy visualization that covers all relevant aspects of privacy, empowering users with the information they require to make well-informed decisions on data disclosure.

8.4 Open questions, limitations and future research

Although a variety of studies were conducted, questions—including those arising from these studies—still remain. Open questions and suggestions for future research will be summarized in the following section. Furthermore, next to the limitations discussed in the previous chapters for each individual study, general limitations of this dissertation will be addressed in this section.

First, one problem in the understanding of online privacy is that it is often operationalized as an umbrella term and approached on a general level instead of being considered in relation to the context and of individual characteristics. Inappropriate research methods and instruments also potentiate the misunderstanding of information privacy (Baek, 2014; Dienlin & Trepte, 2014; Xu et al., 2010). Especially within the privacy paradox literature, the bulk of the research relies upon analyzing perceived intention and not actual behavior. Regardless, the discrepancies between indicated general privacy concerns and actual information disclosure online have been unequivocally established. Despite this, whether a concept as fluid as privacy can be covered with conventional surveys remains questionable, as stated privacy concerns on a general level are almost always different than the actual information disclosure behavior. Measuring actual behavior is a difficult endeavor—especially in the context of privacy—as observing the behavior of participants can be perceived as a privacy intrusion in itself. The experiment presented in Chapter 3 and the Q-sort study discussed in Chapter 5 constitute novel attempts to shed more light on the multifaceted nature of information privacy. However, the context of these two studies can also be considered artificial. Therefore, conducting longitudinal studies with users of online services, observing real behavior and shifting the focus to more qualitative research might prove viable solutions to the problem.

Second, cognitive dissonance theory (Festinger, 1957) suggests that individuals have an inner drive to keep attitudes and beliefs in harmony and avoid dissonance. According to this theory, if dissonance takes place (e.g., not knowing what happens with personal data disclosed online), this emotional state is perceived as uncomfortable, motivating people to achieve consonance (e.g., considering permissions). Furthermore, if people are exposed to dissonant situations, they will avoid further situations that might increase the dissonance (e.g., no longer downloading apps). However, according to the privacy paradox approach, the majority of users are in a dissonant state of perceived insecurity because of concerns regarding new technologies such as smartphones and apps (Benenson et al., 2012; Shklosvski et al., 2014). It is unclear whether users try to achieve consonance through consideration of permissions, by seeking

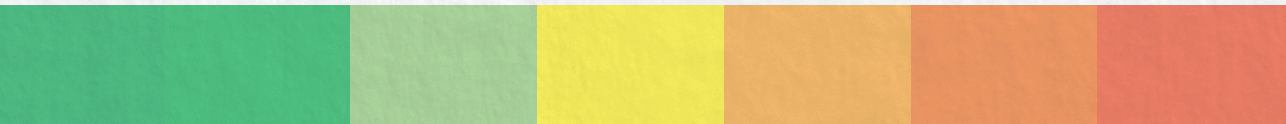
information about how their data are used, or by shifting their boundaries with respect to privacy (e.g., downplaying or ignoring privacy threats). As privacy is influenced by attitudes, operant conditioning could also function as an approach for resolving dissonant states. For instance, every time a user posts information in a social network, this post is liked by others in the network. Repeated rewards in the form of likes strengthen the positive attitude toward posting private information (Skinner, 1938). Receiving rewards can then be considered a positive reinforcement. However, operant conditioning can also result in negative reinforcement. In an online environment, privacy intrusion is intangible, and the consequences are subtle. From the findings of the experiment and interviews, we know that this is also true for users with technical knowledge. The findings obtained in this dissertation strongly indicate that users operate in a dissonant state. However, the question of how users deal with this dissonance remains unanswered and opens possibilities for future studies.

Third, some researchers argue that privacy boundaries are consciously defined and goal-driven (Petronio, 2002; Stanton, 2003). Communication privacy management theory suggests that users form virtual privacy spaces defined by boundaries. These privacy boundaries play a crucial role in the decision-making process of whether to disclose information and with whom (Petronio 2002, 2016). The violation of these boundaries by other parties might be perceived as harmful, depending on the risk-control assessment. In this sense, users might evaluate information disclosure as unacceptable and might consequently act upon this consideration, for instance, not downloading an app because of privacy concerns or moving boundaries with regard to privacy concerns (Sutanto et al., 2013). Although a rational weighing of risks against benefits is unlikely, the decision of whether to use an online service seems to be indeed conscious to a large extent. Of course, consciousness does not prevent internal or external influences from introducing bias. Future research could therefore focus on the formation of privacy boundaries and the reasoning behind pushing privacy boundaries back and forth to explain the privacy paradox. Although information privacy cannot be a one-size-fits-all solution, entirely relying on individual characteristics is also problematic when satisfying the functional complexity of online privacy. Therefore, a starting point in such research could be clustering privacy perceptions into groups of users, as shown in the Q-sort study. However, the focus of this study is still limited, as it pertains to the context and the technology itself. As such, to gain a deeper understanding of the processes explaining seemingly paradoxical online behavior, deeper insight into the aspects users' value most and the severity of privacy intrusion needs to be achieved.

Fourth, findings on the role of technical knowledge and privacy sensitivity are diverse. Where some researchers found a positive effect on privacy protection attributable to familiarity with the technical aspects of online services (Ion et al., 2015; Ketelaar & Van Balen, 2018), others could not confirm this relation (De Luca et al., 2016; Reidenberg et al., 2015). Research with experts in the realm of this dissertation showed that technical knowledge does not automatically protect against the potential dangers of cognitive biases and heuristic thinking. Nevertheless, it is likely that while digital literacy plays a role in privacy behavior, being digitally literate is more than being familiar with the technical aspects of online services, as was approached in the realm of this dissertation. Y. J. Park (2011) suggests that awareness about institutional practices and an understanding of privacy policy play a role in disclosure behavior. Therefore, in agreement with Y. J. Park (2011), promoting digital literacy among all kinds of users (privacy fundamentalists, privacy pragmatists and unconcerned users) is needed. Although privacy is a multifaceted concept, communicating privacy must strive toward an integrated approach that covers as many privacy facets as possible without overwhelming the user with too much information.

8.5 Conclusion

When we talk about information privacy, the axiom of a democratic society is that individuals who are part of this society can decide for themselves the extent personal information is shared, disclosed or sold to third parties, for which purpose and the retention period. In an online context, data subjects would arguably act differently if they knew what the actual value of their personal data is or that their personal data would be shared with other parties a priori to the data exchange (Acquisti et al, 2017; Varian, 2002). The current data economy is largely characterized by power imbalances between the data subject and data holder. This is a relationship destined to fail, and steps must be taken toward improved privacy through legal and social contracts that are governed by the principles of notice and choice. It is therefore important to continue the public, political, legal, and scientific discourse on privacy. In this discourse, the protection of the right of each individual to decide autonomously where their privacy boundaries lie should be central. Users of online services must have the right to decide *what* information they are willing to disclose, to *whom* and in *which* particular situation. This decision must be embedded within a given societal structure and associated norms, the technology itself and supported by law and regulations (A. L. Allen, 1999; Nissenbaum, 2004, 2011; Solove 2002).



References

- Abdul-Ghani, H. A., & Dimitri K. (2019). A comprehensive study of security and privacy guidelines, threats, and countermeasures: An IoT perspective. *Journal of Sensor and Actuator Networks*, 8(22), 1-38. <https://doi.org/10.3390/jsan8020022>
- Abric, J.-C. (1996). Specific processes of social representations. *Papers on Social Representations*, 5(1), 77-80. <http://psr.iscte-iul.pt/index.php/PSR/article/view/211>
- Ackerman M., S., Cranor L. F., & Reagle, J. (1999). Privacy in e-commerce: Examining user scenarios and privacy preferences. *Proceedings of the 1st ACM Conference on Electronic Commerce, USA*, 1-8. <https://doi.org/10.1145/336992.336995>
- Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. *Proceedings of the 5th ACM Conference on Electronic Commerce, USA*, 21-29. <https://doi.org/10.1145/988772.988777>
- Acquisti, A. (2009). Nudging privacy. The behavioral economics of personal information. *IEEE Security & Privacy*, 7(6), 82-85. <https://doi.org/10.1109/MSP.2009.163>
- Acquisti, A. (2010). *The economics of personal data and the economics of privacy*. Carnegie Mellon University.
- Acquisti, A., & Fong, C. (2020). An experiment in hiring discrimination via online social networks. *Management Science* 66(3), 1005-1024. <https://doi.org/10.1287/mnsc.2018.3269>
- Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 3(1), 26-33. <https://doi.org/10.1109/MSP.2005.22>
- Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., & Wilson, S. (2017). Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys*, 50(3), 1-41. <https://doi.org/10.1145/3054926>
- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509-514. <https://doi.org/10.1126/science.aaa1465>
- Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, 54(2), 442-492. <https://doi.org/10.1257/jel.54.2.442>
- Aditya, P., Bhattacharjee, B., Druschel, P., Erdélyi, V., & Lentz, M. (2014). Brave new world: Privacy risks for mobile users. *Proceedings of the ACM MobiCom workshop on Security and privacy in mobile environments, USA*, 7-12. <http://dx.doi.org/10.1145/2646584.2646585>
- Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, 24(4), 665-694. <https://doi.org/10.2307/3250951>
- Agarwal, R., Sambamurthy, V., & Stair, R. (1997). Cognitive absorption and the adoption of new information technologies. In L. Dosier, & J. Keys (Eds.), *Academy of Management Best Paper Proceedings* (pp. 293-297). <https://doi.org/10.5465/ambpp.1997.4983719>

- Ahmadian, A. S., Strüber, D., & Jürjens, J. (2019). Privacy-enhanced system design modeling based on privacy features. *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, USA*, 1492-1499. <https://doi.org/10.1145/3297280.3297431>
- AICPA and CICA. (2009). *Generally accepted privacy principles* (Reports 0001069). The American Institute of Certified Public Accountants, Inc. and The Canadian Institute of Chartered Accountants. [http://op.bna.com.s3.amazonaws.com/pl.nsf/id/byul-7xhufa/\\$File/gapp.pdf](http://op.bna.com.s3.amazonaws.com/pl.nsf/id/byul-7xhufa/$File/gapp.pdf)
- Ajzen, I. (1985). From intentions to actions: a theory of planned behavior. In J. Kuhl, & J. Beckman (Eds.), *Action-Control: From cognition to behavior* (pp. 11-39). Springer-Verlag.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.
- Akdeniz, Y. (2002). Anonymity, democracy, and cyberspace. *Social Research*, 69(1), 180-194. <https://www.jstor.org/stable/40971545>
- Alashoor, T., & Baskerville, R. (2015). *The privacy paradox: The role of cognitive absorption in the social networking activity*. Paper presented at the Thirty Sixth International Conference on Information Systems, Fort Worth, Texas, USA, 1-20. <https://aisel.aisnet.org/icis2015/proceedings/SecurityIS/5/>
- Alimadadi, A., Aryal, S., Manandhar, I., Munroe, P. B., Joe, B., & Cheng, X. (2020). Artificial intelligence and machine learning to fight COVID-19. *Physiological Genomics*, 52(4), 200-202. <https://doi.org/10.1152/physiolgenomics.00029.2020>
- Al-Jamal, M., & Abu-Shanab, E. (2015). Privacy policy of e-government websites: An itemized checklist proposed and tested. *Management Research Practice*, 7(3), 80-95. <https://doi.org/10.15849/icit.2015.0066>
- Allen, A. L. (1999). Coercing privacy. *William & Mary Law Review*, 40(3), article 3, 723-757. <https://scholarship.law.wm.edu/wmlr/vol40/iss3/3>
- Allen, A. L. (2012). What must we hide: The ethics of privacy and the ethos of disclosure. *St. Thomas Law Review*, 25, 1-19. https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1556&context=faculty_scholarship
- Allen, A. L. (2016). Protecting one's own privacy in a big data economy. *Harvard Law Review Forum*, 130, 71-78. <https://harvardlawreview.org/2016/12/protecting-ones-own-privacy-in-a-big-data-economy/>
- Allen, C. (2019, October 28). *How decades of data privacy law birthed GDPR*. <https://rubica.com/brief-history-of-consumer-data-privacy-law/>
- Alsaleh, M., Alomar, N., & Alarifi, A. (2017). Smartphone users: Understanding how security mechanisms are perceived and new persuasive methods. *PLoS ONE*, 12(3), art. e0173284. <https://doi.org/10.1371/journal.pone.0173284>

- Alshammari, M. & Simpson, A. (2017). Towards a principled approach for engineering privacy by design. In E. Schweighofer, H. Leitold, A. Mittrakas, & K. Rannenber (Eds.). *Lecture Notes in Computer Science (including subseries Security and Cryptology)*: Vol. 10518. *Privacy Technologies and Policy* (pp. 161-177). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-67280-9_9
- Altman, I. (1975). *The environment and social behavior*. Brooks/Cole.
- Ambrose, M. L., & Ausloos, J. (2013). The right to be forgotten across the pond. *Journal of Information Policy*, 3, 1-23. <https://doi.org/10.5325/jinfopoli.3.2013.0001>
- Androulidakis, I., & Kandus, G. (2011). Mobile phone security awareness and practices of students in Budapest. *Proceedings of the Sixth International Conference on Digital Telecommunications, Hungary*, 1, 118-124.
- Anthony, S. F. (2001, July, 1). *The case for standardization of privacy policy formats*. Federal Trade Commission. <https://www.ftc.gov/public-statements/2001/07/case-standardization-privacy-policy-formats>
- Antón, A. I., Bertino, E., Li, N., & Yu, T. (2007). A roadmap for comprehensive online privacy policy management. *Communications of the ACM*, 50(7), 109-116. <https://doi.org/10.1145/1272516.1272522>
- Antón, A. I., Earp, J. B., He, Q., Stufflebeam, W., Bolchini, D., & Jensen, C. (2004). Financial privacy policies and the need for standardization. *IEEE Security & Privacy*, 2(2), 36-45. <https://doi.org/10.1109/MSECP.2004.1281243>
- Anwar, M. J., Gill, A. Q., & Beydoun, G. (2018). A review of information privacy laws and standards for secure digital ecosystems. In Australasian Conference on Information Systems (Ed.), *Australasian Conference on Information Systems 2018*, 36 (pp. 1-12). University of Technology Sydney ePress. <https://doi.org/10.5130/acis2018.bb>
- App Innovation. (2015). List & Notes (version 2.6.15). Google Play Store. <https://play.google.com/store/apps/details?id=com.ListAndNote.gen&gl=GB>
- Apthorpe, N., Shvartzshnaider, Y., Mathur, A., Reisman, D., & Feamster, N. (2018). Discovering smart home internet of things privacy norms using contextual integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2), 1-23. <https://doi.org/10.1145/3214262>
- Arkko, J. (2020) The influence of internet architecture on centralised versus distributed internet services, *Journal of Cyber Policy*, 5(1), 30-45. <https://doi.org/10.1080/23738871.2020.1740753>
- Asia-Pacific Economic Cooperation. (2005). *APEC privacy framework* (Reports APEC#2005-SO-01.2). Asia Pacific Economic Cooperation Secretariat.
- Ausloos, J. (2012). The 'right to be forgotten'– worth remembering? *Computer Law & Security Review*, 28(2), 143-152. <https://doi.org/10.1016/j.clsr.2012.01.006>

- Auxier, B., Rainie, L., Anderson, M., Perrin, A., Kumar, M., & Turner, E. (2019, November, 15). *Americans and privacy: Concerned, confused and feeling lack of control over their personal information* (Report). Pew Research Center.
- Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M., & Kalhori, S. R. N. (2020). Predicting Covid-19 incidence through analysis of Google trends data in Iran: Data mining and deep learning pilot study. *JMIR Public Health and Surveillance*, 6(2): e18828m, 1-6. <https://doi.org/10.2196/18828>
- Baek, Y. M. (2014). Solving the privacy paradox: A counter-argument experimental approach. *Computers in Human Behavior*, 38, 33-42. <https://doi.org/10.1016/j.chb.2014.05.006>
- Bal, G. (2014). Designing privacy indicators for smartphone app markets: A new perspective on the nature of privacy risks of apps. *Proceedings of the 20th Americas Conference on Information Systems, USA*, 1-11. <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1206&context=amcis2014&httpsredir=1&referer=>
- Baldassarre, M. T., Barletta, V. S., Caivano, D., & Scalera, M. (2019). Privacy oriented software development. In M. Piattini, P. Rupino da Cunha, I. García Rodríguez de Guzmán, & R. Pérez-Castillo (Eds.), *Quality of Information and Communications Technology: Vol. 1010. Communications in Computer and Information Science* (pp. 18-32). https://doi.org/10.1007/978-3-030-29238-6_2
- Balebako, R., Schaub, F., Adjerid, I., Acquisti, A., & Cranor, L. F. (2015). The impact of timing on the salience of smartphone app privacy notices. *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices, USA*, 63-74. <https://doi.org/10.1145/2808117.2808119>
- Bandara, R. (2020). *The ethics of online privacy in the data-driven marketplace: A power-responsibility equilibrium and construal level theory perspective* (Doctoral dissertation, School of Management, Operations and Marketing, University of Wollongong, Dubai). <https://ro.uow.edu.au/theses1/738>
- Bandara, R., Fernando, M., & Akter, S. (2017). The privacy paradox in the data-driven marketplace. The role of knowledge deficiency and psychological distance. *Procedia Computer Science*, 121, 562-567. <https://doi.org/10.1016/j.procs.2017.11.074>
- Bandara, R., Fernando, M., & Akter, S. (2019). Addressing privacy predicaments in the digital marketplace: A power-relations perspective. *International Journal of Consumer Studies*, 44(5), 423-434. <https://doi.org/10.1111/ijcs.12576>
- Bandura, A. (1969). *Principles of behavior modification*. Holt, Rinehart & Winston.
- Bansal, G., Zahedi, F. M., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support System*, 49(2), 138-150. <https://doi.org/10.1016/j.dss.2010.01.010>

- Barati, M. & Rana, O. (2020). Enhancing user privacy in IoT: Integration of GDPR and blockchain. In Z. Zheng, H.-N. Dai, M. Tang, & X. Chen (Eds.), *Communications in Computer and Information Science: Vol. 1156. Blockchain and Trustworthy Systems* (pp. 322–335). https://doi.org/10.1007/978-981-15-2777-7_26
- Barbas, S. (2012). Saving privacy from history. *DePaul Law Review*, 61(4), 1-77. <https://ssrn.com/abstract=2155320>
- Barkhuus, L. (2012). The mismeasurement of privacy: Using contextual integrity to reconsider privacy in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 367-376. <https://doi.org/10.1145/2207676.2207727>
- Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, 11(9). <https://doi.org/10.5210/fm.v11i9.1394>
- Barth, S., & De Jong, M. D. T. (2017). The privacy paradox - Investigating discrepancies between expressed privacy concerns and actual online behavior - A systematic literature review. *Telematics and Informatics*, 34(7), 1038–1058. <https://doi.org/10.1016/j.tele.2017.04.013>
- Barth, S., De Jong, M. D. T., & Junger, M. (2020). *Lost in privacy? Online privacy from a cybersecurity expert perspective*. Manuscript submitted for publication.
- Barth, S., De Jong, M. D. T., Junger, M., Hartel P. H., & Roppelt, J. C. (2019). Putting the privacy paradox to the test. Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. *Telematics and Informatics*, 41, 55-69. <https://doi.org/10.1016/j.tele.2019.03.003>
- Barth, S., Ionita, D., De Jong, M. D. T., Hartel, P. H., & Junger, M. (in press). Privacy Rating: A user-centered approach for visualizing data handling practices of online services. *IEEE Transactions on Professional Communication*.
- Barth, S., Ngo, T., De Jong, M. D. T., & Krämer, N. C. (2020). *Toward an understanding of online privacy perceptions: Using the Q-sort method to identify different user perspectives*. Manuscript submitted for publication.
- Beldad, A., De Jong, M. D. T., & Steehouder, M. (2010). Reading the least read? Indicators of users' intention to consult privacy statements on municipal websites. *Government Information Quarterly*, 27(3), 238-244. <https://doi.org/10.1016/j.giq.2010.01.004>
- Belli, L., Schwartz, M., & Louzada, L. (2017). Selling your soul while negotiating the conditions: from notice and consent to data control by design. *Health and Technology*, 7, 453–467. <https://doi.org/10.1007/s12553-017-0185-3>
- Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. (2004). International differences in information privacy concerns: A global survey of consumers. *The Information Society*, 20(5), 313–324. <https://doi.org/10.1080/01972240490507956>

- Benenson, Z., Kroll-Peters, O., & Krupp, M. (2012). Attitudes to IT security when using a smartphone. *Proceedings of the Federated Conference on Computer Science and Information System, Poland*, 1179-1183. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6354359>
- Ben-Shahar, O., & Chilton, A. (2016). Simplification of privacy disclosures: An experimental test. *The Journal of Legal Studies*, 45(S2), S41-S67. <https://doi.org/10.1086/688405>
- Benton, K., Camp, L. J., & Garg, V. (2013). Studying the effectiveness of Android application permissions requests. *Proceedings of the 2013 IEEE International Conference on Pervasive Computing and Communications Workshops, USA*, 291-296. <https://doi.org/10.1109/PerComW.2013.6529497>
- Berendt, B., Günther, O., & Spiekermann, S. (2005). Privacy in e-commerce: stated preferences vs. actual behaviour. *Communications of the ACM*, 48(4), 101-106. <https://doi.org/10.1145/1053291.1053295>
- Beresford, A. R., Rice, A., Skehin, N., & Sohan, R. (2011). MockDroid: trading privacy for application functionality on smartphones. *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications, USA*, 49-54. <https://doi.org/10.1145/2184489.2184500>
- Bergström, A. (2015). Online privacy concerns: a broad approach to understanding the concerns of different groups for different uses. *Computers in Human Behavior*, 53, 419-426. <https://doi.org/10.1016/j.chb.2015.07.025>
- Berners-Lee, T., Handler, J., & Lassila, O. (2001). The semantic Web. *Scientific American*, 284(5), 34-43. <https://www.jstor.org/stable/26059207>
- Beznosov, K., Inglesant, P., Lobo, J., Reeder, R., & Zurko, M. E. (2009). Usability meets access control: challenges and research opportunities. *Proceedings of the 14th ACM symposium on Access control models and technologies, USA*, 73-74. <https://doi.org/10.1145/1542207.1542220>
- Bier, C., Birnstill, P., Krempel, E., Vagts, H., & Beyerer, J. (2014). Enhancing privacy by design from a developer's perspective. In B. Preneel, & D. Ikonomou (Eds.), *Lecture Notes in Computer Science: Vol. 8319. Privacy Technologies and Policy*. (pp. 73-85). https://doi.org/10.1007/978-3-642-54069-1_5
- Bincoletto, G. (2019). A data protection by design model for privacy management in electronic health records. In M. Naldi, G. F. Italiano, K. Rannenber, M. Medina, & A. Bourka (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11498. Privacy Technologies and Policy* (pp. 161-181). https://doi.org/10.1007/978-3-030-21752-5_11
- Blumer, H. (1986). *Symbolic interactionism: perspectives and method*. University of California Press.
- Blumler, J. G., & Katz, E. (1974). *The uses of mass communications: current perspectives on gratifications research*. Sage Beverly Hills.

- Boanabeau, E. (2014). Introduction. In M. Finneran Dennedy, J. Fox, & T. R. Finneran (Eds.), *The Privacy Engineer's Manifesto. Getting from Policy to Code to QA to Value* (pp. xxxv-xxxvi). Apress. <https://doi.org/10.1007/978-1-4302-6356-2>
- Boehme-Neßler, V. (2016). Privacy: a matter of democracy. Why democracy needs privacy and data protection. *International Data Privacy Law*, 6(3), 220-229. <https://doi.org/10.1093/idpl/ipw007>
- Boritz, J. E., & No, W. G. (2011). E-commerce and privacy: Exploring what we know and opportunities for future discovery. *Journal of Information Systems*, 25(2), 11-45. <https://doi.org/10.2308/isys-10090>
- Branscomb, A. W. (1994). *Who Owns Information? From privacy to public access*. BasicBooks.
- Branscomb, A. W. (1995). Anonymity, autonomy, and accountability: Challenges to the First Amendment in cyberspaces. *The Yale Law Journal*, 104(7), 1639-1679. <https://doi.org/10.2307/797027>
- Bräunlich K., Dienlin, T., Eichenhofer, J., Helm, P., Trepte, S., Grimm, R., Seubert, S., & Gusy, C. (2020). Linking loose ends: An interdisciplinary privacy and communication model. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/1461444820905045>
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Brenton, M. (1964). *The privacy invaders*. Coward McCann.
- Brown, I. (2014). Britain's smart meter programme: A case study in privacy by design. *International Review of Law, Computers & Technology*, 28(2), 172-184. <https://doi.org/10.1080/13600869.2013.801580>
- Brown, S. R. (1986). Q technique and method: Principle and procedures. In W. D. Berry, & M. S. Lewis-Beck (Eds.), *New tools for social scientists: Advances and applications in research methods* (pp. 57-76). Sage.
- Brown, S. R. (1993). A primer on Q methodology. *Operant Subjectivity*, 16(3/4), 91-138. [dx.doi.org/10.15133/j.os.1993.002](https://doi.org/10.15133/j.os.1993.002)
- Bu, F., Wang, N., Jiang, B., & Liang, H. (2020). „Privacy by Design” implementation: Information system engineers' perspective. *International Journal of Information Management*, 53, art. 102124. <https://doi.org/10.1016/j.ijinfomgt.2020.102124>
- Buchmann, J. (2012). *Internet privacy. A multidisciplinary analysis*. acatech - Deutsche Akademie der Technikwissenschaften.
- Buck, C., Horbel, C., Germelmann, C. C., & Eymann, T. (2014). The unconscious app consumer: Discovering and comparing the information-seeking patterns among mobile application consumers. *Proceedings of the Twenty Second European Conference on Information Systems, Israel*, 1-14. <https://aisel.aisnet.org/ecis2014/proceedings/track14/8/>

- Buck, C., Horbel, C., Kessler, T., & Christian, C. (2014). Mobile consumer apps: Big data brother is watching you. *Marketing Review St. Gallen*, 31, 26-34. <https://doi.org/10.1365/s11621-014-0318-2>
- Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision making. *Journal of Mathematical Psychology*, 50(3), 220–241. <https://doi.org/10.1016/j.jmp.2006.01.003>
- Bygrave, L. A. (2017). Data protection by design and by default: Deciphering the EU's legislative requirements. *Oslo Law Review*, 2(4), 105-120. <https://doi.org/10.18261/issn.2387-3299-2017-02-03>
- California Consumer Privacy Act of 2018 [1798.100 - 1798.199.100] (Title 1.81.5 added by Stats. 2018, Ch. 55, Sec. 3.) (2020).
- Canadian Standards Association. CSA Group. (2014). *Model code for the protection of personal information of 23 December 2014, first published in March 1996; reaffirmed 2001* (CAN/CSA-Q830-96). https://www.afn.ca/uploads/files/nihbforum/info_and_privacy_doc_csa_model_code_for_the_protection_of_personal_information.pdf
- Casey, R. L. (2016). *Privacy by Design: Setting a new standard for privacy certification* (Reports 15-2971-H). Deloitte LLP.
- Castelluccia, C. (2012). Behavioural tracking on the Internet: A technical perspective. In S. Gutwirth, R. Leenes, P. de Hert, & Y. Pouillet (Eds.), *European data protection: In good health?* (pp. 21-33). Springer Science+Business Media B.V. https://doi.org/10.1007/978-94-007-2903-2_2
- Cate, F. H. (1997). *Privacy in the information age*. Brookings Institution Press.
- Cavoukian, A. (2006). *Creation of a global privacy standard*. Information and Privacy Commissioner of Ontario. http://www.ehcca.com/presentations/privacysymposium1/cavoukian_2b_h5.pdf
- Cavoukian, A. (2009). *Privacy by design. The 7 foundational principles*. Information and Privacy Commissioner of Ontario. <http://dataprotection.industries/wp-content/uploads/2017/10/privacy-by-design.pdf>
- Cavoukian, A. (2010). *Privacy by design. The 7 foundational principles* (Technical report). Information and Privacy Commissioner of Ontario, Ontario, Canada.
- Cavoukian, A. (2011a). Privacy by design: Best practices for privacy and the smart grid. In N. Pohlmann, H. Reimer, & Wolfgang Schneider (Eds.), *ISSE 2010 Securing Electronic Business Processes* (pp. 260–270). Vieweg+Teubner. https://doi.org/10.1007/978-3-8348-9788-6_25
- Cavoukian, A. (2011b). *Privacy by Design. The 7 Foundational Principles* (Technical report, revised version). Information and Privacy Commissioner of Ontario. https://iapp.org/media/pdf/resource_center/pbd_implement_7found_principles.pdf
- Cavoukian, A. (2012). Privacy by design [leading edge]. *IEEE Technology and Society Magazine* 31(4), 18-19. <https://doi.org/10.1109/MTS.2012.2225459>

- Cavoukian, A. (2020). Understanding how to implement privacy by design, one step at a time. *IEEE Consumer Electronics Magazine*, 9(2), 78-82. <https://doi.org/10.1109/MCE.2019.2953739>
- Cavoukian, A., & Chibba, M. (2016). Cognitive cities, big data and citizen participation: The essentials of privacy and security. In E. Portmann, & M. Finger (Eds.), *Studies in Systems, Decision and Control: Vol. 63. Towards Cognitive Cities* (pp. 61–82). https://doi.org/10.1007/978-3-319-33798-2_4
- Chai, S., Bagchi-Sen, S. Morrell, C., Rao, H. R., & Upadhyaya, S. J. (2009). Internet and online information privacy: An exploratory study of preteens and early teens. *IEEE Transaction on Professional Communication*, 52(2), 167-182. <https://doi.org/10.1109/TPC.2009.2017985>
- Chen, H.-T., & Chen, W. (2015). Couldn't or wouldn't? The influence of privacy concerns and self-efficacy in privacy management on privacy protection. *Cyberpsychology, Behavior, and Social Networking*, 18(1), 13-19. <https://doi.org/10.1089/cyber.2014.0456>
- Chen, Y., Zha, M., Zhang, N., Xu, D., Zhao, Q., Feng, X., Yuan, K., Suya, F., Tian, Y., Chen, K., Wang, X. F., & Zou, W. (2019). Demystifying hidden privacy settings in mobile apps. *Proceedings of the 2019 IEEE Symposium on Security and Privacy, USA*, 850–866. <https://doi.org/10.1109/SP.2019.00054>
- Chia, P. H., Yamamoto, Y., & Asokan, N. (2012). Is this app safe? A large scale study on application permissions and risk signals. *Proceedings of the 21st International Conference on World Wide Web, France*, 311–320. <https://doi.org/10.1145/2187836.2187879>
- Chin, E., Felt, A. P., Sekar, V., & Wagner, D. (2012). Measuring user confidence in smartphone security and privacy. *Proceedings of the 8th Symposium on Usable Privacy and Security, USA*, 1-16. <https://doi.org/10.1145/2335356.2335358>
- Choi, H., Park, J., & Jung, Y. (2018). The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior*, 81, 42-51. <https://doi.org/10.1016/j.chb.2017.12.001>
- Colesky, M. & Caiza, J. C. (2018). A system of privacy patterns for informing users: Creating a pattern system. *Proceedings of the 23rd European Conference on Pattern Languages of Programs, USA*, Article 16, 1-11. <https://doi.org/10.1145/3282308.3282325>
- Colesky, M., Caiza, J. C., Del Álamo, J. M., Hoepman, J.-H., & Martin, Y.-S. (2018). A system of privacy patterns for user control. *Proceedings of the 33rd Annual ACM Symposium on Applied Computing, USA*, 1150–1156. <https://doi.org/10.1145/3167132.3167257>
- Colesky, M., Hoepman, J.-H., & Hillen, C. (2016). A critical analysis of privacy design strategies. *Proceedings of the 2016 IEEE Symposium on Security and Privacy Workshops, USA*, 33–40. <https://doi.org/10.1109/SPW.2016.23>
- Colonna, L. (2013). Prism and the European Union's Data Protection Directive. *The John Marshall Journal of Information Technology & Privacy Law*, 30(2), article 1, 1-27. <http://repository.jmls.edu/jitpl/vol30/iss2/1>

- Com2uS. (2016). Tower Defense: Infinite War (version 1.2.4). Google Play Store. <https://play.google.com/store/apps/details?id=com.com2us.towerdefenseplus.normal2.freefull.google.global.android.common&gl=GB>
- Commission Decision of 26 July 2000 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the safe harbour privacy principles and related frequently asked questions issued by the US Department of Commerce (notified under document number C(2000) 2441). Legislation *OJ L 215*. The European Parliament and the Council of European Union. (2000/520/EC). Publication Office of the European Union.
- Conti, G., & Sobiesk, E. (2010). Malicious interface design: exploiting the user. *Proceedings of the 19th international conference on World wide web, USA*, 271-280. <https://doi.org/10.1145/1772690.1772719>
- Craig, T., & Ludloff, M. E. (2011). *Privacy and big data: The players, regulators, and stakeholders*. O'Reilly Media, Inc.
- Cranor, L. F. (2009). *Find web sites that respect your privacy*. CMU Usable Privacy and Security Laboratory, Carnegie Mellon University. <http://www.privacybird.org>
- Crazelle Solutions. (2015). My ToDo List. Google Play Store. Abandoned.
- Crossler, R. E., & Bélanger, F. (2017). The mobile privacy-security knowledge gap model: understanding behaviors. *Proceedings of the 50th Hawaii International Conference on System Sciences, USA*, 4071-4080. <https://doi.org/10.24251/hicss.2017.491>
- Crutchfield, R. S. (1955). Conformity and character. *American Psychologist*, 10(5), 191-198. <https://doi.org/10.1037/h0040237>
- Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science*, 10(1), 340-347. <https://doi.org/10.1287/orsc.10.1.104>
- Davies, N. & Langheinrich, M. (2013). Privacy by design [From the editor in chief]. *IEEE Pervasive Computing*, 12(2), 2-4. <https://doi.org/10.1109/MPRV.2013.34>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Davis, K., Frederick, W. C., & Blomstrom, R. L. (1980). *Business and society: Concepts and policy issues*. New York, NY: McGraw-Hill.
- Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly*, 47(1), 1-15. <https://doi.org/10.1086/268763>
- De Jong, M. D. T., & Wu, Y. (2018). Functional complexity and Web site design: Evaluating the online presence of UNESCO World Heritage sites. *Journal of Business and Technical Communication*, 32(3), 347-372. <https://doi.org/10.1177/1050651918762029>

- De Jong, S., & Spagnuolo, D. (2020). Iconified representations of privacy policies: A GDPR perspective. In Á. Rocha, H. Adeli, L. Reis, S. Costanzo, I. Orovic, & F. Moreira (Series Eds.), *Trends and Innovations in Information Systems and Technologies: Vol. 2. Advances in intelligent systems and computing* (Vol. 1160, pp. 796-806). Springer. https://doi.org/10.1007/978-3-030-45691-7_75
- De Lima Salgado, A., Silva Dias, F., Rodrigues Mattos, J. P., Pontin de Mattos Fortes, R., & Hung, P. C. K. (2019). Smart toys and children's privacy: usable privacy policy insights from a card sorting experiment. *Proceedings of the 37th ACM International Conference on the Design of Communication, USA*, Article 16, 1-8. <https://doi.org/10.1145/3328020.3353951>
- De Luca, A., Das, S., Ortlieb, M., Ion, I., & Laurie, B. (2016). Expert and non-expert attitudes towards (secure) instant messaging. *Proceeding of the Twelfth Symposium on Usable Privacy and Security, USA*, 147-157. <https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-de-luca.pdf>
- De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313(5787), 684-687. <https://doi.org/10.1126/science.1128356>
- Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009). Facebook and online privacy: attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, 15(1), 83-108. <https://doi.org/10.1111/j.1083-6101.2009.01494.x>
- Dehling, T., Gao, F., Schneider, S., & Sunyaev, A. (2015). Exploring the far side of mobile health: Information security and privacy of mobile health apps on iOS and Android. *JMIR mHealth uHealth*, 3(1), e8, 1-17. <https://doi.org/10.2196/mhealth.3672>
- Deonn Games ltd. (2013). Astroid Defense Classic. <https://play.google.com/store/apps/details?id=com.deonn.games.ad2&gl=GB>
- Department for Digital Culture Media & Sport. (2018). *Secure by Design: Improving the cyber security of consumer Internet of Things report* (Secure by Design Report). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/775559/Secure_by_Design_Report_.pdf
- Deuker, A. (2010). Addressing the privacy paradox by expanded privacy awareness - the example of context-aware services. In M. Bezzi, P. Duquenoy, S. Fischer-Hüber, M. Hansen, & G. Zhang (Series Eds.), *Privacy and identity management for life. Privacy and identity 2009. IFIP advances in information and communication technology* (Vol. 320, pp. 275-283). Springer. https://doi.org/10.1007/978-3-642-14282-6_23
- Dienlin, T., & Trepte, S. (2014). Is the privacy paradox a relic of the past? *An in-depth analysis of privacy attitudes and privacy behaviors. European Journal of Social Psychology*, 45(3), 285-297. <https://doi.org/10.1002/ejsp.2049>
- Digital Advertising Alliance. (2016). *Put the YourAdChoices icon to work for you*. YourAdChoices. <http://youradchoices.com/learn>

- Dinev, T. (2014). Why would we care about privacy? *European Journal of Information Systems*, 23(2), 97-102. <https://doi.org/10.1057/ejis.2014.1>
- Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, 17(1), 61-80. <https://doi.org/10.1287/isre.1060.0080>
- Directive 95/46/EC of the European Parliament and of the protection of individuals with regard to the processing of personal data and on the free movement of such data. Legislation No L 281/31. The European Parliament and the Council of European Union. (1995). Publication Office of the European Union.
- Dogruel, L., Joeckel, S., & Bowman, N. D. (2015). Choosing the right app. An exploratory perspective on heuristic decision processes for smartphone app selection. *Mobile Media & Communication*, 3(1), 125-144. <https://doi.org/10.1177/2050157914557509>
- Donnenwerth, G. V., & Foa, U. G. (1974). Effect of resource class on retaliation to injustice in interpersonal exchange. *Journal of Personality and Social Psychology*, 29(6), 785-793. <https://doi.org/10.1037/h0036201>
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2), 135-150. <https://doi.org/10.1086/257897>
- Drgon, M., Magnuson, G., & Sabo, J. (2016). *Privacy management reference model and methodology (PMRM) Version 1.0. Committee Specification 02 (PMRM-v1.0-cs02. Standards Track Work Product)*. <http://docs.oasis-open.org/pmr/PMRM/v1.0/cs02/PMRM-v1.0-cs02.html>.
- Drozd, O. (2016). Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process. In D. Aspinall, J. Camenisch, M. Hansen, S. Fischer-Hübner, & C. Raab (Series Eds.), *Privacy and Identity Management. Time for a Revolution?. Privacy and Identity. IFIP Advances in Information and Communication Technology* (Vol. 476, pp. 129-140). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-41763-9_9
- Edelman, B. (2009). Adverse selection in online “trust” certifications. *Proceedings of the 11th International Conference on Electronic Commerce, USA*, 205-212. <https://doi.org/10.1145/1593254.1593286>
- Edwards, B., Hofmeyr, S., & Forrest, S. (2016). Hype and heavy tails: A closer look at data breaches. *Journal of Cybersecurity*, 2(1), 3-14. <https://doi.org/10.1093/cybsec/tyw003>
- Edwards, L., & Abel, W. (2014). *The use of privacy icons and standard contract terms for generating consumer trust and confidence in digital services* (CREATE Working Paper Series 2014/15). Zenodo. <https://doi.org/10.5281/zenodo.12506>
- Efroni Z., Metzger, J., Mischau, L., & Schirmbeck, M. (2019). Privacy icons: A risk-based approach to visualisation of data processing. *European Data Protection Law Review*, 5(3), 352-366. <https://doi.org/10.21552/edpl/2019/3/9>

- Egele, M., Kruegel, C., Kirda, E., & Vigna, G. (2011). PiOS: detecting privacy leaks in iOS applications. *Proceedings of the 18th Annual Network and Distributed System Security Symposium, USA*, 77–183. <https://www.ndss-symposium.org/wp-content/uploads/2017/09/egel.pdf>
- Egelman, S., Kannavara, R., & Chow, R. (2015). Is this thing on? Crowdsourcing privacy indicators for ubiquitous sensing platforms. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, USA*, 1669–1678. <https://doi.org/10.1145/2702123.2702251>
- Emami-Naeini, P., Dixon, H., Agarwal, Y., & Cranor, L. F. (2019). Exploring how privacy and security factor into IoT device purchase behavior. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, USA*, Article 534, 1–12. <https://doi.org/10.1145/3290605.3300764>
- Enck, W., Gilbert, P., Han, S., Tendulkar, V., Chun, B.-G., Cox, L. P., Jung, J., McDaniel, P., & Sheth, A. N. (2014). TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems*, 32(2), article 5, 1–15. <https://doi.org/10.1145/2619091>
- Esayas S., Mahler, T., & McGillivray, K. (2016). Is a picture worth a thousand terms? Visualising contract terms and data protection requirements for cloud computing users. In S. Casteleyn, P. Dolog, & C. Pautasso (Series Eds.), *Current Trends in Web Engineering. Lectures Notes in Computer Science* (Vol. 9881, pp. 39–56). Springer. https://doi.org/10.1007/978-3-319-46963-8_4
- Fairweather, J. R. (2001). Factor stability, number of significant loadings and interpretation: evidence from three case studies and suggested guidelines. *Operant Subjectivity*, 25(1), 37–58. [dx.doi.org/10.15133/j.os.2001.012](https://doi.org/10.15133/j.os.2001.012)
- Farnden, J., Martini, B., & Choo, K-K. R. (2015). Privacy risks in mobile dating apps. *Proceedings of 21st Americas Conference on Information Systems, Puerto Rico*, 1–16. <https://arxiv.org/abs/1505.02906v1>
- Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., & Wagner, D. (2012). Android permissions: User attention, comprehension, and behavior. *Proceedings of the Eighth Symposium on Usable Privacy and Security, USA*, article 3, 1–14. <https://doi.org/10.1145/2335356.2335360>
- Fernback, J., & Papacharissi, Z. (2007). Online privacy as legal safeguard: the relationship among consumer, online portal, and privacy policies. *New Media & Society*, 9(5), 715–734. <https://doi.org/10.1177/1461444807080336>
- Festinger L. (1957). *A theory of cognitive dissonance*. Stanford University Press.
- Finneran Denny, M., Fox, J., & Finneran, T. R. (2014). *The Privacy Engineer's Manifesto. Getting from policy to code to QA to value*. Apress. <https://doi.org/10.1007/978-1-4302-6356-2>

- Flender, C., & Müller, G. (2012). Type indeterminacy in privacy decisions: the privacy paradox revisited. In J. Busemeyer, F. Dubois, A. Lambert-Mogiliansky, & M. Melucci (Series Eds.), *Quantum interaction. Lecture notes in Computer Science* (Vol. 7620, pp. 148-159). Springer. https://doi.org/10.1007/978-3-642-35659-9_14
- Foa, U. G. (1971). Interpersonal and economic resources. *Science*, *171*(3969), 345–351. <https://doi.org/10.1126/science.171.3969.345>
- Forbrukerrådet. (2018). *Deceived by Design. How tech companies use dark patterns to discourage us from exercising our rights to privacy* (Report: Deceived by design). <https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf>
- Fox, G., Tonge, C., Lynn, T., & Mooney, J. (2018). Communicating compliance: Developing a GDPR privacy label. *Proceedings of the 24th Americas Conference on Information Systems, USA*, *3*, 1867-1871.
- Franke, G., Clever, T., Van Dijk, W., Raider, J., & De Jonge, R. (2019). Privacy label. Blog series part I-IV. Sensor Lab. <https://medium.com/sensor-lab/the-privacy-illusion-994ed98ec3ab>
- Friedland, G., & Sommer, R. (2010). Cybercasing the joint: on the privacy implications of geo-tagging. *Proceedings of the 5th USENIX conference on Hot topics in security, USA*, 1-6. https://static.usenix.org/events/hotsec10/tech/full_papers/Friedland.pdf
- Gallagher, K., Patil, S., & Memon, N. (2017). New me. Understanding expert and non-expert perceptions and usage of the Tor anonymity network. *Proceedings of the Thirteenth Symposium on Usable Privacy and Security, USA*, 385-398. <https://www.usenix.org/system/files/conference/soups2017/soups2017-gallagher.pdf>
- Gambino, A., Kim., J., Sundar, S. S., Ge, J., & Rosson, M. B. (2016). User disbelief in privacy paradox: Heuristics that determine disclosure. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, USA*, 2837–2843. <https://doi.org/10.1145/2851581.2892413>
- Garg, V., & Camp, J. (2013). Heuristics and biases: Implications for security design. *IEEE Technology and Society magazine*, *32*(1), 73-79. <https://doi.org/10.1109/MTS.2013.2241294>
- Gavison, R. (1992). Too early for a requiem: Warren and Brandeis were right on privacy vs. free speech. *South Carolina Law Review*, *43*(3), 437-471. <https://scholarcommons.sc.edu/sclr/vol43/iss3/3>
- Gellman, R. (2019). *Fair information practices: A basic history*. (Report Version 2.19), 1-51. <http://dx.doi.org/10.2139/ssrn.2415020>
- Gerunov, A. A. (2020). Attitudes towards privacy by design in e-government: Views from the trenches. *Journal of Social and Administrative Science*, *7*(1), 1-17.
- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. University of California Press.

- Gisch, M., De Luca, A., & Blanchebarbe, M. (2007). The privacy badge: A privacy-awareness user interface for small devices. *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology, Singapore*, 583-586. <https://doi.org/10.1145/1378063.1378159>
- Glancy, D. J. (1979). The invention of the right to privacy. *Arizona Law Review*, 21(1), 1-40. <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1318&context=facpubs>
- Gleitman, H., Gross, J., & Reisberg, D. (2011). *Psychology*. W. W. Norton & Company, Inc.
- Gomez, J., Pinnick, T., & Soltani, A. (2009). *KnowPrivacy*. UC Berkeley, School of Information. https://ashkansoltani.files.wordpress.com/2013/01/knowprivacy_final_report.pdf
- González v. Google Spain, C-131/12 (Court of Justice of the European Union 2014).
- Graf, C., Hochleitner, C., Wolkerstorfer, P., Angulo, J., Fischer-Hübner, S., & Wästlund, E. (2011). *PrimeLife – Privacy and Identity Management in Europe for Life* (Final HCI Research Report). Center for Usability Research and Engineering, Karlstads universitet, Unabhängiges Landeszentrum für Datenschutz. http://primelife.ercim.eu/images/stories/deliverables/d4.1.5-final_hci_research_report-public.pdf
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Canada*, paper 534, 1-14. <https://doi.org/10.1145/3173574.3174108>
- Gu, J., Xu, Y. (C.), Xu, H., Zhang, C., & Ling, H. (2017). Privacy concerns for mobile app download: An elaboration likelihood model perspective. *Decision Support Systems*, 94, 19-28. <https://doi.org/10.1016/j.dss.2016.10.002>
- Guidelines for the Regulation of Computerized Personal Data Files. Adopted by General Assembly resolution 45/49 of 14 December 1990. United Nations. (1990).
- Guo, X., Sun, Y., Yan, Z., & Wang, N. (2012). Privacy-personalization paradox in adoption of mobile health service: The mediating role of trust. *Proceedings of the Pacific Asia Conference on Information Systems, Vietnam*, paper 27, 1-16. <https://aisel.aisnet.org/pacis2012/27>
- Gutwirth, S., & Hildebrandt, M. (2008). *Profiling the European citizen*. Cross-disciplinary perspectives. Springer Science + Business Media B.V. <https://doi.org/10.1007/978-1-4020-6914-7>
- Haenlein, M., & Kaplan, A. (2019). A brief history of Artificial Intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5-14. <https://doi.org/10.1177/0008125619864925>
- Hafetz, J. L. (2002). "A man's home is his castle?": Reflections on the home, the family, and privacy during the late nineteenth and early twentieth centuries. *William & Mary Journal of Women and the Law*, 8(2), article 2, 175-242. <https://scholarship.law.wm.edu/wmjowl/vol8/iss2/2>

- Haninger, K., & Thompson, K. M. (2004). Content and ratings of teen-rated video games. *JAMA*, 291(7), 856–865. <https://doi.org/10.1001/jama.291.7.856>
- Hansen, M. (2009). Putting privacy pictograms into practice. A European perspective. In S. Fischer, E. Maehle, & R. Reischuk (Series Eds.), *Informatik 2009 - Im Focus das Leben. Lecture Notes in Informatics* (Vol. P-154, pp. 1703-1716). Gesellschaft für Informatik e.V.. <https://dl.gi.de/bitstream/handle/20.500.12116/31137/194.pdf?sequence=1&isAllowed=y>
- Hansen, M., Schwartz, A., & Cooper, A. (2008). Privacy and identity management. *IEEE Security & Privacy*, 6(2), 38–45. <https://doi.org/10.1109/MSP.2008.41>
- Harsanyi, J. C. (1967). Games with incomplete information played by “Bayesian” players, I–III part I. The basic model. *Management Science*, 14(3), 159-182. <https://doi.org/10.1287/mnsc.14.3.159>
- Hawkey, K., & Inkpen, K. M. (2007). PrivateBits: managing visual privacy in web browsers. *Proceedings of Graphics Interface, USA*, 215–223. <https://doi.org/10.1145/1268517.1268553>
- Hazarika, B., Khuntia, J., Parthasarathy, M., & Karimi, J. (2016). Do hedonic and utilitarian apps differ in consumer appeal? In V. Sugumaran, V. Yoon, & M. Shaw (Series Eds.), *E-Life: Web-Enabled Convergence of Commerce, Work, and Social Life. WEB 2015. Lecture Notes in Business Information Processing* (Vol. 258, pp. 233-237). Springer. https://doi.org/10.1007/978-3-319-45408-5_28
- Heinonen, K., & Pura, M. (2006). Classifying mobile services. *Proceedings of Helsinki Mobility Roundtable. Sprouts: Working Papers on Information Systems*, 6(42), 1-18. <http://sprouts.aisnet.org/6-42>
- Hill, H. (2014). Aus Daten Sinn machen: Analyse- und Deutungskompetenzen in der Datenflut. *Die Öffentliche Verwaltung*, 6(3), 213-222.
- Hoepman, J.-H. (2014). Privacy design strategies (Extended abstract). In N. Cuppens-Bouahia, F. Cuppens, S. Jajodia, A. Abou El Kalam, & T. Sans (Series Eds.), *ICT Systems Security and Privacy Protection. IFIP Advances in Information and Communication Technology* (Vol. 428, pp. 446-459). Springer-Verlag. https://doi.org/10.1007/978-3-642-55415-5_38
- Hoffmann, C. P., Lutz, C., & Ranzini, G. (2016). Privacy cynicism: a new approach to the privacy paradox. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 10(4), article 7, 1-18. <https://doi.org/10.5817/CP2016-4-7>
- Holtz, L.-E., Nocun, K., & Hansen, M. (2011a). Towards displaying privacy information with icons. In S. Fischer-Hübner, P. Duquenoy, M. Hansen, R. Leenes, & G. Zhang (Series Eds.), *Privacy and Identity Management for Life. IFIP Advances in Information and Communication Technology* (Vol. 352, pp. 338-348). Springer. https://doi.org/10.1007/978-3-642-20769-3_27

- Holtz, L.-E., Zwingelberg, H., & Hansen, M. (2011b). Privacy policy icons. In J. Camenisch, S. Fischer-Hübner, & K. Rannenberg (Eds.), *Privacy and Identity Management for Life* (pp. 279-285). Springer. https://doi.org/10.1007/978-3-642-20317-6_15
- Holvast, J. (2009). A history of privacy. In V. Matyáš, S. Fischer-Hübner, D. Cvrček, & P. Švenda (Eds.), *The future of identity in the information society. IFIP Advances in Information and Communication Technology* (Vol. 298, pp. 13-42). Springer. <http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-3-642-03315-5>
- Hornyack, P., Han, S., Jung, J., Schechter, S., & Wetherall, D. (2011). These aren't the droids you're looking for: retrofitting android to protect from imperious applications. *Proceedings for the 18th ACM conference on Computer and communication security, USA*, 639-652. <https://doi.org/10.1145/2046707.2046780>
- Hu, Q., & Ma, S. (2010). Does privacy still matter in the era of Web 2.0? A qualitative study of user behavior towards online social networking activities. *Proceedings of Pacific Asia Conference on Information Systems, Taiwan*, 2, 591-602. <http://www.pacis-net.org/file/2010/S14-03.pdf>
- Huckvale K., Prieto, J. T., Tilney, M., Benghozi, P.-J., & Car, J. (2015). Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Medicine*, 13(1), article 214, 1-13. <https://doi.org/10.1186/s12916-015-0444-y>
- Hughes-Roberts, T. (2012). A cross-disciplined approach to exploring the privacy paradox: explaining disclosure behaviour using the theory of planned behavior. *Proceedings of the UK Academy for Information Systems Conference, United Kingdom*, paper 7, 1-16. <http://aisel.aisnet.org/ukais2012/7>
- Hughes-Roberts, T. (2013). Privacy and social networks: is concern a valid indicator of intention and behaviour? *Proceedings of the 2013 IEEE International Conference on Social Computing, USA*, 909-912. <https://doi.org/10.1109/SocialCom.2013.140>
- Iannella, R., & Finden, A. (2009). Privacy awareness: Icons and expression for social networks. *Proceedings of the 8th International Workshop for Technical, Economic and Legal Aspects of Business Models for Virtual Goods, Incorporating the 6th International Open Digital Rights Language Workshop (ODRL 2009), Belgium*, 1-13. http://virtualgoods.org/2010/VirtualGoodsBook2010_13.pdf
- Ingenious Software solution. (2013). Tasks+ To Do List Manager. Google Play Store. Abandoned.
- Introna, L. D. (1997). Privacy and the computer: Why we need privacy in the information society. *Metaphilosophy*, 28(3), 259-275. <https://doi.org/10.1111/1467-9973.00055>
- Ion, I., Reeder, R., & Consolvo, S. (2015). '... No one can hack my mind.' Comparing expert and non-expert security practices. *Proceedings of Eleventh Symposium on Usable Privacy and Security, Canada*, 327-346. <https://www.usenix.org/system/files/conference/soups2015/soups15-paper-ion.pdf>
- Irwin, F. W. (1953). Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality*, 21(3), 329-335. <https://doi.org/10.1111/j.1467-6494.1953.tb01775.x>

- Itbe.pl. (2015). EveryDay ToDo List Task List. Google Play Store. Abandoned.
- Jensen C., & Potts, C. (2004). Privacy policies as decision-making tools: an evaluation of online privacy notices. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austria*, 471-478. <https://doi.org/10.1145/985692.985752>
- Jia, H., Wisniewski, P., Xu, H., Rosson, M. B., & Carroll, J. M. (2015). Risk-taking as a learning process for shaping teen's online information privacy behaviors. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Canada*, 583-599. <https://doi.org/10.1145/2675133.2675287>
- Joeckel, S., Dogruel, L., & Bowman, N. D. (2017). The reliance on recognition and majority vote heuristics over privacy concerns when selecting smartphone apps among German and US consumers. *Information, Communication & Society*, 20(4), 621-636. <https://doi.org/10.1080/1369118X.2016.1202299>
- Johnson, S. D., Blythe, J. M., Manning, M., & Wong, G. T. W. (2020). The impact of IoT security labelling on consumer product choice and willingness to pay. *PLoS ONE*, 15(1), e0227800, 1–21. <https://doi.org/10.1371/journal.pone.0227800>
- Joinson, A.N., Reips, U.-D., Buchanan, T., & Paine Schofield, C.B. (2010). Privacy, trust, and self-disclosure online. *Human-Computer Interaction*, 25(1), 1–24. <https://doi.org/10.1080/07370020903586662>
- Jorgensen, Z., Chen, J., Gates, C.S., Li, N., Proctor, R.W., & Yu, T. (2015). Dimensions of risk in mobile applications. A user study. *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy, USA*, 49-60. <https://doi.org/10.1145/2699026.2699108>
- Junglas, I. A., Johnson, N. A., & Spitzmüller, C. (2008). Personality traits and concern for privacy: An empirical study in the context of location-based services. *European Journal of Information Systems*, 17(4), 387–402. <https://doi.org/10.1057/ejis.2008.29>
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5), 1449–1475. <https://doi.org/10.1257/000282803322655392>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 263–292. <https://doi.org/10.2307/1914185>
- Kang, R., Dabbish, L., Fruchter, N., & Kiesler, S. (2015). “My data just goes everywhere:” User mental models of the Internet and implications for privacy and security. *Proceedings of 11th Symposium on Usable Privacy and Security, Canada*, 39-52. <https://www.usenix.org/system/files/conference/soups2015/soups15-paper-kang.pdf>
- Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Uses and gratifications research. *The Public Opinion Quarterly*, 37(4), 509–523. <https://www.jstor.org/stable/2747854>
- Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, 25(6), 607-635. <https://doi.org/10.1111/isj.12062>

- Kehr, F., Wentzel, D., & Kowatsch, T. (2014). Privacy paradox revised. Pre-existing attitudes, psychological ownership, and actual disclosure. *Proceedings of the Thirty Fifth International Conference on Information Systems, New Zealand*, 1-12. https://www.alexandria.unisg.ch/242216/1/PDF%20Proof_revised.pdf
- Keith, M. J., Thompson, S. C., Hale, J., Lowry, P. B., & Greer, C. (2013). Information disclosure on mobile devices: R-examining privacy calculus with actual user behavior. *International Journal of Human-Computer Studies*, 71(12), 1163–1173. <https://doi.org/10.1016/j.ijhcs.2013.08.016>
- Kelbert, F., Shirazi, F., Simo, H., Wüchner, T., Buchmann, J., Pretschner, A., & Waidner, M. (2012). State of online privacy: A technical perspective. In J. Buchmann (Series Ed.), *Internet Privacy. A multidisciplinary analysis. acatech STUDY* (Vol. September 2012, pp. 189-279). Springer Vieweg. https://doi.org/10.1007/978-3-642-31943-3_4
- Kelley P. G., Bresee, J., Cranor, L. F., & Reeder, R. W. (2009). A 'nutrition label' for privacy. *Proceedings of the 5th Symposium on Usable Privacy and Security, USA*, article 4, 1-12. <https://doi.org/10.1145/1572532.1572538>, art. 4.
- Kelley P. G., Cesca, L., Bresee, J., & Cranor, L. F. (2010). Standardizing privacy notices: An online study of the nutrition label approach. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, USA*, 1573-1582. <https://doi.org/10.1145/1753326.1753561>
- Kelley, P. G., Consolvo, S., Cranor, L. F., Jung, J., Sadeh, N., & Wetherall, D. (2012). A conundrum of permissions: Installing applications on an Android smartphone. In J. Blyth, S. Dietrich, & L. J. Camp (Series Eds.), *Financial Cryptography and Data security. Lecture Notes in Computer Science* (Vol. 7398, pp. 68-79). Springer. https://doi.org/10.1007/978-3-642-34638-5_6
- Kelley, P. G., Cranor, L. F., & Sadeh, N. (2013). Privacy as part of the app decision-making process. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, France*, 3393–3402. <https://doi.org/10.1145/2470654.2466466>
- Kerr, O. S. (2004). The Fourth Amendment and new technologies: Constitutional myths and the case for caution. *Michigan Law Review*, 102(5), 801-888. <https://doi.org/10.2307/4141982>
- Ketelaar, P. E., & Van Balen, M. (2018). The smartphone as your follower. The role of smartphone literacy in the relation between privacy concerns, attitude and behaviour towards phone-embedded tracking. *Computers in Human Behavior*, 78, 174-182. <https://doi.org/10.1016/j.chb.2017.09.034>
- Khatoon, A., & Corcoran, P. (2017). Android permission system and user privacy - A review of concept and approaches. *Proceedings of 2017 IEEE 7th International Conference on Consumer Electronics-Berlin, Germany*, 153-158. <https://doi.org/10.1109/ICCE-Berlin.2017.8210616>

- Kim, G. S., Park, S.-B., & Oh, J. (2008). An examination of factors influencing consumer adoption of short message service (SMS). *Psychology & Marketing*, 25(8), 769-786. <https://doi.org/10.1002/mar.20238>
- Klitou, D. (2012). A solution, but not a panacea for defending privacy: The challenges, criticism and limitations of privacy by design. In B. Preneel & D. Ikononou (Series Eds.), *Privacy Technologies and Policy. Annual Privacy Forum. Lecture Notes in Computer Science* (Vol. 8319, pp. 86-110). Springer. https://doi.org/10.1007/978-3-642-54069-1_6
- Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 64, 122-134. <https://doi.org/10.1016/j.cose.2015.07.002>
- Kooappa LLC. (2013). Safe the cave: Tower Defense. Google Play Store. Abandoned.
- Kool L., Timmer, J., Royakkers, L., & Van Est, R. (2017). *Urgent Upgrade. Protect public values in our digitized society*. Rathenau Instituut. https://www.rathenau.nl/sites/default/files/2018-03/Urgent_Upgrade.pdf
- Koops B.-J., & Leenes, R. (2014). Privacy regulation cannot be hardcoded. A critical comment on the 'privacy by design' provision in data-protection law. *International Review of Law, Computers & Technology*, 28(2), 159-171. <https://doi.org/10.1080/13600869.2013.801589>
- Korzaan, M., Brooks, N., & Greer, T. (2009). Demystifying personality and privacy: An empirical investigation into antecedents of concerns for information privacy. *Journal of Behavioral Studies in Business*, 1, 1-17. <https://www.aabri.com/manuscripts/09225.pdf>
- Kozyreva, A., Herzog, S., Lorenz-Spreen, P., Hertwig, R., & Lewandowsky, S. (2020). *Artificial intelligence in online environments: Representative survey of public attitudes in Germany*. Max Planck Institute for Human Development. <https://doi.org/10.17617/2.3188061>
- Krämer, N. C., & Haferkamp, N. (2011). Online self-presentation: balancing privacy concerns and impression construction on social networking sites. In S. Trepte & L. Reinecke (Eds.), *Privacy Online* (pp. 127-141). Springer. https://doi.org/10.1007/978-3-642-21521-6_10
- Kramp, T., Van Kranenburg, R., & Lange, S. (2013). Introduction to the Internet of Things. In A. Bassi, M. Bauer, M. Fiedler, T. Kramp, R. Van Kranenburg, S. Lange, & S. Meissner (Eds.), *Enabling Things to Talk. Designing IoT solutions with the IoT Architectural Reference Model* (pp. 1-10). Springer. <https://doi.org/10.1007/978-3-642-40403-0>
- Kraus, L., Wechsung, I., & Möller, S. (2014). A comparison of privacy and security knowledge and privacy concern as influencing factors for mobile protection behavior. In A. Morton (Chair), *Behavior*. Workshop conducted at the symposium on Usable Privacy and Security, Canada. <https://cups.cs.cmu.edu/soups/2014/workshops/privacy/s2p4.pdf>

- Kucuk, S. U. (2009). The evolution of market equalization on the Internet. *Journal of Research for Consumers*, 16, 1-15. http://jrconsumers.com/Academic_Articles/issue_16/market_equalization_paper_academic.pdf
- Kucuk, S. U. (2016). Consumerism in the digital age. *The Journal of Consumer Affairs*, 50(3), 515-538. <https://doi.org/10.1111/joca.12101>
- Kumaraguru, P., & Cranor, L. F. (2005). *Privacy indexes: A survey of Westin's studies*. (Report No. CMU-ISRI-5-138). Institute for Software Research International, School of Computer Science, Carnegie Mellon University. <https://www.cs.cmu.edu/~ponguru/CMU-ISRI-05-138.pdf>
- Kunreuther, H. (1984). Causes of underinsurance against natural disasters. *The Geneva Papers on Risk and Insurance – Issues and Practice*, 9, 206–220. <https://doi.org/10.1057/gpp.1984.12>
- Kupfer, J. (1987). Privacy, autonomy, and self-concept. *American Philosophical Quarterly*, 24(1), 81-89. <https://www.jstor.org/stable/20014176>
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2), 443–478. <https://doi.org/10.1162/003355397555253>
- Langheinrich, M. (2001). Privacy by design - Principles of privacy-aware ubiquitous systems. In G. D. Abowd, B. Brumitt, & S. Shafer (Series Eds.), *Ubicomp 2001: Ubiquitous Computing. Lecture Notes in Computer Science* (Vol. 2201, pp. 273-291), Springer. https://doi.org/10.1007/3-540-45427-6_23
- LaRose, R., & Rifon, N. J. (2007). Promoting i-safety: Effects of privacy warnings and privacy seals on risk assessment and online privacy behavior. *The Journal of Consumer Affairs*, 41(1), 127-149. <https://doi.org/10.1111/j.1745-6606.2006.00071.x>
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: a literature review and two-component model. *Psychological Bulletin*, 107(1), 34–47. <https://doi.org/10.1037/0033-2909.107.1.34>
- Lentz, L., & Pander Maat, H. (2004). Functional analysis for document design. *Technical Communication*, 51(3), 387-398. <https://www.jstor.org/stable/43090556>
- Levy, S. E., & Gutwin, C. (2005). Improving understanding of website privacy policies with fine-grained policy anchors. *Proceedings of the 14th international conference on World Wide Web, USA*, 480–488. <https://doi.org/10.1145/1060745.1060816>
- Li, H., Sarathy, R., & Xu, H. (2010). Understanding situational online information disclosure as a privacy calculus. *Journal of Computer Information Systems*, 51(1), 62–71. <https://doi.org/10.1080/08874417.2010.11645450>
- Li, Y. (2012). Theories in online information privacy research: a critical review and an integrated framework. *Decision Support Systems*, 54(1), 471–481. <https://doi.org/10.1016/j.dss.2012.06.010>

- Liccardi, I., Pato, J., Weitzner, D. J., Abelson, H., & De Roure, D. (2014). No technical understanding required: Helping users make informed choices about access to their personal data. *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, United Kingdom*, 140-150. <https://doi.org/10.4108/icst.mobiquitous.2014.258066>
- Loewenstein, G. (1999). Because it is there: The challenge of mountaineering...for utility theory. *Kyklos*, 52(3), 315-344. <https://doi.org/10.1111/j.1467-6435.1999.tb00221.x>
- Lowdermilk, T. (2013). *User-centered design: a developer's guide to building user-friendly applications*. O'Reilly Media, Inc.
- Lutz, C., & Strathoff, P. (2011). Privacy concerns and online behavior – not so paradoxical after all? Viewing the privacy paradox through different theoretical lenses. In S. Brändli, R. Schister, & A. Tamò (Eds.), *Multinationale Unternehmen und Institutionen im Wandel – Herausforderungen für Wirtschaft, Recht und Gesellschaft* (pp. 81-99). Stämpfli Verlag. <http://dx.doi.org/10.2139/ssrn.2425132>
- Lwin, M., Wirtz, J., & Williams, J. D. (2007). Consumer online privacy concerns and responses: a power-responsibility equilibrium perspective. *Journal of the Academy of Marketing Science*, 35, 572-585. <https://doi.org/10.1007/s11747-006-0003-3>
- Lyon, D. (2003). *Surveillance as social sorting*. New York, NY: Routledge.
- Makri, E.-L., & Lambrinouidakis, C. (2015). Privacy principles: Towards a common privacy audit methodology. In S. Fischer-Hübner, C. Lambrinouidakis, & J. López (Series Eds.), *Trust, privacy and security in digital business. Lecture Notes in Computer Science (including subseries Security and Cryptology)* (Vol. 9264, pp. 219–234). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-22906-5_17
- Manheim, K., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy. *Yale Journal of Law & Technology*, 21, 106-188. <https://ssrn.com/abstract=3273016>
- Manier, M. J., & O'Brien Louch, M. (2010). Online social networks and the privacy paradox: a research framework. *Issues in Information Systems*, XI(1), 513-517. https://doi.org/10.48009/1_iis_2010_513-517
- Martin, J., & Christin, N. (2016). Ethics in cryptomarket research. *International Journal of Drug Policy*, 35, 84–91. <https://doi.org/10.1016/j.drugpo.2016.05.006>
- Martin, K. (2013). Transaction costs, privacy, and trust: The laudable goals and ultimate failure of notice and choice to respect privacy online. *First Monday*, 18(12). <https://doi.org/10.5210/fm.v18i12.4838>
- Martin, K. (2016a). Understanding privacy online: Development of a social contract approach to privacy. *Journal of Business Ethics*, 137(3), 551–569. <https://doi.org/10.1007/s10551-015-2565-9>
- Martin, K. (2016b). Data aggregators, consumer data, and responsibility online: Who is tracking consumers online and should they stop? *The Information Society*, 32(1), 51-63. <https://doi.org/10.1080/01972243.2015.1107166>

- Martin, K., & Shilton, K. (2016). Putting mobile application privacy in context: An empirical study of user privacy expectations for mobile devices. *The Information Society*, 32(3), 200-216. <https://doi.org/10.1080/01972243.2016.1153012>
- Martín, Y.-S., & Del Álamo, J. M. (2017). A metamodel for privacy engineering methods. In J. M. del Álamo, S. F. Gürses, & A. Datta (Series Eds.), *Proceedings of the 3rd International Workshop on Privacy Engineering co-located with 38th IEEE Symposium on Security and Privacy, CEUR Workshop Proceedings* (Vol. 1873, pp.41-48). http://ceur-ws.org/Vol-1873/IWPE17_paper_24.pdf
- Martín, Y.-S., Del Alamo, J. M., & Yelmo, J. C. (2014). Engineering privacy requirements valuable lessons from another realm. *Proceedings of 2014 IEEE 1st International Workshop on Evolving Security and Privacy Requirements Engineering, Karlskrona*, 19–24. <https://doi.org/10.1109/ESPRES.2014.6890523>
- Marx, G. T. & Muschert, G. W. (2007). Personal information, borders, and the new surveillance studies. *Annual Review of Law and Social Science*, 3, 375-395. <https://doi.org/10.1146/annurev.lawsocsci.3.081806.112824>
- Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Finding from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction, USA*, article 81, 1-32. <https://doi.org/10.1145/3359183>
- Matz, S. C., Appel, R. E., & Kosinski, M. (2020). Privacy in the age of psychological targeting. *Current Opinion in Psychology*, 31, 116-121. <https://doi.org/10.1016/j.copsyc.2019.08.010>
- Mayfield, K. (2016). Pseudonymisation: a 20-yearold idea never seemed so timely. *Journal of Direct, Data Digital Marketing Practices*, 17, 222–226. <https://doi.org/10.1057/s41263-016-0005-x>
- Mehldau, M. (2007). *Iconset for data-privacy declarations vo.1*. <https://netzpolitik.org/wp-upload/data-privacy-icons-vo1.pdf>.
- Metzger, M. J. (2007). Communication privacy management in electronic commerce. *Journal of Computer-Mediated Communication*, 12(2), 335–361. <https://doi.org/10.1111/j.1083-6101.2007.00328.x>
- Meynhardt, T. (2009). Public value inside: what is public value creation? *International Journal of Public Administration*, 32(3-4), 192–219. <https://doi.org/10.1080/01900690902732632>
- Michelfelder, D. P. (2001). The moral value of informational privacy in cyberspace. *Ethics and Information Technology*, 3, 129-135. <https://doi.org/10.1023/A:1011802227136>
- Mihalich DS Group. (2013). ToDo list - Private Tasks (version 1.8). Google Play Store. <https://play.google.com/store/apps/details?id=app.mds.privatetasks&gl=GB>

- Miller, A. R. (1969). Personal privacy in the computer age: The challenge of a new technology in an information-oriented society. *Michigan Law Review*, 67(6), 1089-1246. <https://doi.org/10.2307/1287516>
- Milne G. R., Rohm, A. J., & Bahl, S. (2004). Consumers' protection of online privacy and identity. *The Journal of Consumer Affairs*, 38(2), 217-232. <https://doi.org/10.1111/j.1745-6606.2004.tb00865.x>
- Morales-Trujillo, M. E., Matla-Cruz, E. O., García-Mireles, G. A., & Piattini, M. (2018). A systematic mapping study on Privacy by Design in software engineering. *CLEI Electronic Journal*, 22(1), paper 4, 1-29. <https://doi.org/10.19153/cleiej.22.1.4>
- Morel, V., & Pardo, R. (2020). SoK: Three facets of privacy policies. *Proceedings of the 19th Workshop on Privacy in the Electronic Society, Virtual Event, USA*, 41-56. <https://doi.org/10.1145/3411497.3420216>
- Morosan, C., & DeFranco, A. (2015). Disclosing personal information via hotel apps: a privacy calculus perspective. *International Journal of Hospitality Management*, 47, 120-130. <https://doi.org/10.1016/j.ijhm.2015.03.008>
- Moscovici, S. (1984). The phenomenon of social representations. In R. M. Farr, & S. Moscovici, S. (Eds.), *Social Representations* (pp. 3-69). University Press.
- Moskowitz, B., & Raskin, A. (2011, updated June, 28). Privacy Icons. In *Mozilla Wiki*. https://wiki.mozilla.org/Privacy_Icons
- Motiwala, L. F., Li, X., & Liu, X. (2014). Privacy paradox: Does stated privacy concerns translate into the valuation of personal information?. *Proceeding of the 19th Pacific Asia Conference on Information Systems, China*, paper 281. <https://aisel.aisnet.org/pacis2014/281>
- Motti, V. G., & Caine, K. (2016). Towards a visual vocabulary for privacy concepts. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 1078-1082. <https://doi.org/10.1177/1541931213601249>
- Mourey, J. A., & Waldman, A. E. (2020). Past the privacy paradox: The importance of privacy changes as a function of control and complexity. *Journal of the Association for Consumer Research*, 5(2), 162-180. <https://doi.org/10.1086/708034>
- Müller, G., Flender, C., & Peters, M. (2012). Vertrauensinfrastruktur und Privatheit als ökonomische Fragestellung. In J. Buchmann (Series Ed.), *Internet Privacy. A multidisciplinary analysis. acatech STUDY* (Vol. September 2012, pp. 143-188). Springer Vieweg. https://doi.org/10.1007/978-3-642-31943-3_3
- Myers West, S. (2019). Data capitalism: Redefining the logics of surveillance and privacy. *Business & Society*, 58(1), 20-41. <https://doi.org/10.1177/0007650317718185>
- Nagy, J., & Pecho, P. (2009). Social networks security. *Proceedings of the Third International Conference on Emerging Security Information, Systems and Technologies (SecureWare), Greece*, 740-746. <https://doi.org/10.1109/SECURWARE.2009.56>

- National Conference of State Legislatures. (2020, June 6). *Privacy Protections in State Constitutions*. <https://www.ncsl.org/research/telecommunications-and-information-technology/privacy-protections-in-state-constitutions.aspx>
- Newman, A. L. (2015). What the “right to be forgotten” means for privacy in a digital age. *Science*, 347(6221), 507–508. <https://doi.org/10.1126/science.aaa4603>
- Nikiforakis, N., Kapravelos, A., Joosen, W., Kruegel, C., Piessens, F., & Vigna, G. (2013). Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. *Proceedings of the 2013 IEEE Symposium on Security and Privacy, USA*, 541-555. <https://doi.org/10.1109/SP.2013.43>
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 119–158. <https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10>
- Nissenbaum, H. (2010). *Privacy in Context. Technology, policy, and the integrity of social life*. Stanford University Press.
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32-48. https://doi.org/10.1162/DAED_a_00113
- Nissenbaum, H. (2015). Respecting context to protect privacy: Why meaning matters. *Science and Engineering Ethics*, 24, 831-852. <https://doi.org/10.1007/s11948-015-9674-9>
- Nissenbaum, H. (2018). Respecting context to protect privacy: Why meaning matters. *Science and Engineering Ethics*, 24, 831-852. <https://doi.org/10.1007/s11948-015-9674-9>
- Nissenbaum, H. (2019). Contextual integrity up and down the data food chain. *Theoretical Inquiries in Law*, 20(1), 221-256. <https://doi.org/10.1515/til-2019-0008>
- Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *The Journal of Consumer Affairs*, 41(1), 100–126. <https://doi.org/10.1111/j.1745-6606.2006.00070.x>
- Novak, T. P., & Hoffman, D. L. (2008). The fit of thinking style and situation: new measures of situation-specific experiential and rational cognition. *Journal of Consumer Research*, 36(1), 56-72. <https://doi.org/10.1086/596026>
- O’Donoghue, T., & Rabin, M. (2001). Choice and procrastination. *The Quarterly Journal of Economics*, 116(1), 121-160. <https://doi.org/10.1162/003355301556365>
- Ochs, C., & Löw, M. (2012). Un/faire Informationspraktiken: Internet Privacy aus sozialwissenschaftlicher Perspektive. In J. Buchmann (Series Ed.), *Internet Privacy. A multidisciplinary analysis. acatech STUDY* (Vol. September 2012, pp. 15-62). Springer Vieweg.
- Odum, A. L. (2011). Delay discounting: Trait variable? *Behavioural Processes*, 87(1), 1-9. <https://doi.org/10.1016/j.beproc.2011.02.007>
- OECD. (2013). *The OECD privacy framework*. Organisation for Economic Co-operation and Development. https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

- Oetzel, M. C., & Gonja, T. (2011). The online privacy paradox: A social representations perspective. *Proceedings of the CHI Conference on Human Factors in Computing Systems. Extended Abstracts on Human Factors in Computing Systems, Canada*, 2107-2112. <https://doi.org/10.1145/1979742.1979887>
- Office of the Australian Information Commissioner. (2014). *Australian privacy principles. A summary for APP entities from 12 March 2014*. Australian Government. <https://www.oaic.gov.au/assets/privacy/guidance-and-advice/app-quick-reference-tool.pdf>
- Oltramari, A., Piraviperumal, D., Schaub, F., Wilson, S., Cherivirala, S., Norton, T. B., Russell, N. C., Story, P., Reidenberg, J., & Sadeh, N. (2018). PrivOnto: A semantic framework for the analysis of privacy policies. *Semantic Web*, 9(2), 185-203. <https://doi.org/10.3233/SW-170283>
- Oomen, I., & Leenes, R. (2008). Privacy risk perceptions and privacy protection strategies. In E. De Leeuw, S. Fischer-Hübner, J. Tseng, & J. Borking (Eds.), *Policies and Research in Identity Management* (pp. 121-138). Springer. https://doi.org/10.1007/978-0-387-77996-6_10
- Osatuyi, B. (2015). Personality traits and information privacy concern on social media platforms. *Journal of Computer Information Systems*, 55(4), 11-19. <https://doi.org/10.1080/08874417.2015.11645782>
- Palen, L., & Dourish, P. (2003). Unpacking “privacy” for a networked world. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, USA*, 129-136. <https://doi.org/10.1145/642611.642635>
- Pangrazio, L. & Selwyn, N. (2019). “Personal data literacies’: A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*, 21(2), 419-437. <https://doi.org/10.1177/1461444818799523>
- Papacharissi, Z., & Fernback, J. (2005). Online privacy and consumer protection: An analysis of portal privacy statements. *Journal of Broadcasting & Electronic Media*, 49(3), 259-281. https://doi.org/10.1207/s15506878jobem4903_1
- Park, Y. J. (2011). Digital literacy and privacy behavior online. *Communication Research*, 40(2), 215-236. <https://doi.org/10.1177/0093650211418338>
- Park, Y., Ju, J., & Ahn, J.-H. (2015). Are people really concerned about their privacy?: Privacy paradox in mobile environment. *Proceedings of the Fifteenth International Conference on Electronic Business, Hong Kong*, 63, 123-128. <https://aisel.aisnet.org/iceb2015/63>
- Passera, S. (2012). Enhancing contract usability and user experience through visualization. *Proceedings of the 16th International Conference on Information Visualization, France*, 376-382. <https://doi.org/10.1109/IV.2012.69>
- Pavlou, P. A. (2011). State of the information privacy literature: Where are we now and where should we go? *MIS Quarterly*, 35(4), 977-988. <https://doi.org/10.2307/41409969>
- Pavlov, I. (1927). *Conditioned reflexes*. Oxford University Press.

- Pedarsani, P., & Grossglauser, M. (2011). On the privacy of anonymized networks. *Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining, USA*, 1235-1243. <https://doi.org/10.1145/2020408.2020596>
- Pedraza, J., Patricio, M. A., De Asís, A., & Molina, J. M. (2011). Regulatory model for AAL. In E. Corchado, V. Snášel, J. Sedano, A. E. Hassanien, J. L. Calvo, & D. Ślezak (Series Eds.), *Soft Computing Models in Industrial and Environmental Applications, 6th International Conference SOCO 2011, Advances in Intelligent and Soft Computing* (Vol. 87, pp.183-192). Springer. https://doi.org/10.1007/978-3-642-19644-7_20
- Pedraza, J., Patricio, M. A., De Asís, A., & Molina, J. M. (2013). Privacy-by-design rules in face recognition system. *Neurocomputing*, 109, 49-55. <https://doi.org/10.1016/j.neucom.2012.03.023>
- Pennekamp, J., Henze, M., & Wehrle, K. (2017). A survey on the evolution of privacy enforcement on smartphones and the road ahead. *Pervasive and Mobile Computing*, 42, 58-76. <http://dx.doi.org/10.1016/j.pmcj.2017.09.005>
- Pentina, I., Zhang, L., Bata, H., & Chen, Y. (2016). Exploring the privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison. *Computers in Human Behavior*, 65, 409-419. <https://doi.org/10.1016/j.chb.2016.09.005>
- Perera, C., Barhamgi, M., Bandara, A. K., Ajmal, M., Price, B., & Nuseibeh, B. (2020). Designing privacy-aware internet of things applications. *Information Science*, 512, 238-257. <https://doi.org/10.1016/j.ins.2019.09.061>
- Perera, C., McCormick, C., Bandara, A. K., Price, B. A., & Nuseibeh, B. (2016). Privacy-by-Design framework for assessing Internet of Things applications and platforms. *Proceedings of the 6th International Conference on the Internet of Things, Germany*, 83-92. <https://doi.org/10.1145/2991561.2991566>
- Perry, R. (2019). GDPR – project or permanent reality? *Computer Fraud & Security*, 2019(1), 9-11. [https://doi.org/10.1016/S1361-3723\(19\)30007-7](https://doi.org/10.1016/S1361-3723(19)30007-7)
- Personal Information Protection and Electronic Documents Act (PIPEDA), Act S.C. 200, c.5. Current to July 28, 2020, last amended on June 21, 2019 (2000). Office of the Privacy Commissioner of Canada. <https://laws-lois.justice.gc.ca/eng/acts/P-8.6/>
- Peter Semayne v. Richard Gresham, 5 Coke Rep. 91a (1604).
- Petronio, S. (1991). Communication boundary management: a theoretical model of managing disclosure of private information between married couples. *Communication Theory*, 1(4), 311-335. <https://doi.org/10.1111/j.1468-2885.1991.tb00023.x>
- Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. University of New York Press.
- Petronio, S. (2016). Communication privacy management theory. In K. B. Jensen, R. T. Craig, J. D. Pooley, & E. W. Rothenbuhler (Eds.), *The International Encyclopedia of Communication Theory and Philosophy* (pp. 278-286). Wiley-Blackwell. <https://doi.org/10.1002/9781118766804.wbiect138>

- Petterson J. S. (2015). A brief evaluation of icons in the first reading of the European Parliament on COM (2012) 0011. In J. Camenisch, S. Fischer-Hübner, & M. Hansen (Series Eds.), *Privacy and Identity Management for the Future Internet in the Age of Globalisation. IFIP Advances in Information and Communication Technology* (Vol. 457, pp. 125-135). Springer. https://doi.org/10.1007/978-3-319-18621-4_9
- Phelan, C., Lampe, C., & Resnick, P. (2016). It's creepy, but it doesn't bother me. *Proceedings of the CHI Conference on Human Factors in Computing Systems, USA*, 5240-5251. <https://doi.org/10.1145/2858036.2858381>
- Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy Marketing*, 19(1), 27-41. <https://doi.org/10.1509/jppm.19.1.27.16941>
- Pinkas, D. (2016). An eID mechanism built along privacy by design principles using secure elements, pseudonyms and attributes. In D. Hühnlein, H. Roßnagel, C. H. Schunck, & M. Talamo (Series Eds.), *Open Identity Summit 2006, Lecture Notes in Informatics* (Vol. P-264, pp. 93-104). Gesellschaft für Informatik e.V.. <https://subs.emis.de/LNI/Proceedings/Proceedings264/P-264.pdf>
- Pinnick, T. (2011, February, 17). *Privacy short notice design*. TrustArc Inc. <https://www.trustarc.com/blog/?p=1253>
- Poikela, M., Schmidt, R., Wechsung, I., & Möller, S. (2015). FlashPolling privacy: the discrepancy of intention and action in location-based poll participation. *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, Japan*, 813-818. <https://doi.org/10.1145/2800835.2804402>
- Polartouch.se. (2017). New Eskimo Defense. Google Play Store. Abandoned.
- Poneres, K., Hamidi, F., Massey, A., & Hurst, A. (2018). Using icons to communicate privacy characteristics of adaptive assistive technologies. *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Ireland*, 388-390. <https://doi.org/10.1145/3234695.3241003>
- Pötzsch, S. (2009). Privacy awareness: a means to solve the privacy paradox? In V. Matyáš, S. Fischer-Hübner, D. Cvrček, & P. Švenda (Series Eds.), *The Future of Identity in the Information Society. IFIP Advances in Information and Communication Technology* (Vol. 298, pp. 226-236). Springer. https://doi.org/10.1007/978-3-642-03315-5_17
- Pötzsch, S., Wolkerstorfer, P., & Graf, C. (2010). Privacy-awareness information for web forums: results from an empirical study. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, Iceland*, 363-372. <https://doi.org/10.1145/1868914.1868957>

- Preuveneers, D., Joosen, W., & Ilie-Zudor, E. (2016). Data protection compliance regulations and implications for smart factories of the future. *Proceedings of the 12th International Conference on Intelligent Environments, United Kingdom*, 40-47. <https://doi.org/10.1109/IE.2016.15>
- Prichard, J., & Mentzer, K. (2017). An analysis of app privacy statements. *Issues in Information Systems*, 18(4), 179-188. https://doi.org/10.48009/4_iis_2017_179-188
- Privacy Act of 1974, Pub. L. 93-579, § 552a, 88 Stat. 1896 (1974).
- Privacy Label. (2020). *Upgrade your privacy statement: Privacy Label helps you communicate how you use data*. <https://www.privacylabel.org>
- Proctor, R. W., Ali, M. A., & Vu, K.-P. L. (2008). Examining usability of web privacy policies. *International Journal of Human-Computer Interaction*, 24(3), 307-328. <https://doi.org/10.1080/10447310801937999>
- Quay-de la Vallee, H., Selby, P., & Krishnamurthi, S. (2016). On a (per)mission: Building privacy into the app marketplace. *Proceedings of the 6th Workshop on Security and Privacy in Smartphones and Mobile Devices, Austria*, 63-72. <https://doi.org/10.1145/2994459.2994466>
- Quinn, K. (2016). Why we share: a uses and gratifications approach to privacy regulation in social media use. *Journal of Broadcasting & Electronic Media*, 60(1), 61-86. <https://doi.org/10.1080/08838151.2015.1127245>
- Rachovitsa, A. (2016). Engineering and lawyering privacy by design: understanding online privacy both as a technical and international human rights issue. *International Journal of Law and Information Technology*, 24(4), 374-399. <https://doi.org/10.1093/ijlit/eaw012>
- Ramokapane, K. M., Mazeli, A. Z., & Rashid, A. (2019). Skip, skip, skip, accept!!!: A study on the usability of smartphone manufacturer provided default features and user privacy. *Proceedings on Privacy Enhancing Technologies*, 2019(2), 209-227. <https://doi.org/10.2478/popets-2019-0027>
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Legislation OJ L 119, 2.5. The European Parliament and the Council of European Union (2016). Publications Office of the European Union.
- Reidenberg, J. R., Breaux, T., Cranor, L. F., French, B., Grannis, A., Graves, J. T., Liu, F., McDonald, A., Norton, T., Ramanath, R., Russell, N. C., Sadeh, N., & Schaub, F. (2015). Disagreeable privacy policies. Mismatches between meaning and users' understanding. *Berkeley Technology Law Journal*, 30, 1-39. http://ir.lawnet.fordham.edu/faculty_scholarship/619
- Renaud, K., & Shepherd, L. A. (2018). How to make privacy policies both GDPR-compliant and usable. *Proceedings of the International Conference On Cyber Situational Awareness, Data Analytics And Assessment, United Kingdom*, 1-8. <https://doi.org/10.1109/CyberSA.2018.8551442>

- Renaud, K., Volkamer, M., & Renkema-Padmos, A. (2014). Why doesn't Jane protect her privacy? In E. De Cristofaro, & S. J. Murdoch (Series Eds.), *Privacy enhancing technologies. Lecture Notes in Computer Science* (Vol. 8555, pp. 244-262). Springer. https://doi.org/10.1007/978-3-319-08506-7_13
- Rezgui, A., Bouguettaya, A., & Eltoweissy, M. Y. (2003). Privacy on the Web: facts, challenges, and solutions. *IEEE Security & Privacy*, 1(6), 40-49. <https://doi.org/10.1109/MSECP.2003.1253567>
- Ringmann, S. D., Langweg, H., & Waldvogel, M. (2018). Requirements for legally compliant software based on the GDPR. In H. Panetto, C. Debruyne, H. A. Proper, C. A., Ardagna, D. Roman, & R. Meersman (Series Eds.), *On the move to meaningful Internet systems. Lecture Notes in Computer Science (including subseries Programming and Software Engineering)* (Vol. 11230, pp. 258-276). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-02671-4_15
- Roessler, B. (2005). *The value of privacy*. Polity Press.
- Romanou, A. (2018). The necessity of the implementation of privacy by design in sectors where data protection concerns arise. *Computer Law & Security Review*, 34(1), 99-110. <https://doi.org/10.1016/j.clsr.2017.05.021>
- Rossi, A., & Palmirani, M. (2017). A visualization approach for adaptive consent in the European data protection framework. *Proceedings of the 7th International Conference for E-Democracy and Open Government, Austria*, 159-170. <https://doi.org/10.1109/CeDEM.2017.23>
- Rossi, A., & Palmirani, M. (2019). DAPIS: An ontology-based data protection icon set. In G. Peruginelli & S. Faro (Eds.), *Knowledge of the law in the big data age* (pp. 181-195). IOS Press BV. <https://doi.org/10.3233/FAIA190020>
- Rostama, G., Bekhradi, A., & Yannou, B. (2017). From privacy by design to design for privacy. *Proceedings of the 21st International Conference on Engineering Design, Canada*, DS87-6, 317-326. <https://hal.archives-ouvertes.fr/hal-01673578>
- Rothfeder, J. (1992). *Privacy for sale: How computerization has made everyone's private life an open secret*. Simon & Schuster Trade.
- Rothstein, M. A., & Tovino, S. A. (2019). California takes the lead on data privacy law. *Hastings Center Report* 49(5), 4-5. <https://doi.org/10.1002/hast.1042>
- Rouvroy A., & Pouillet Y. (2009). The right to informational self-determination and the value of self-development: Reassessing the importance of privacy for democracy. In S. Gutwirth, Y. Pouillet, P. De Hert, C. de Terwangne, & S. Nouwt (Eds.), *Reinventing data protection?* (pp. 45-76). Springer. https://doi.org/10.1007/978-1-4020-9498-9_2
- Rubin, A.M. (1984). Ritualized and instrumental television viewing. *Journal of Communication*, 34(3), 67-77. <https://doi.org/10.1111/j.1460-2466.1984.tb02174.x>

- Rudolph M., Feth, D., & Polst, S. (2018). Why users ignore privacy policies - A survey and intention model for explaining user privacy behavior. In M. Kurosu (Series Ed.), *Human-Computer Interaction: Theories, Methods, and Human Issues. Lecture Notes in Computer Science* (Vol. 10901, pp. 587-598). Springer. https://doi.org/10.1007/978-3-319-91238-7_45
- Sabo, J. T. (2007). ISTPA Operational analysis of international privacy requirements. In N. Pohlmann, H. Reimer, & W. Schneider (Eds.), *ISSE/SECURE 2007 Securing Electronic Business Processes* (pp. 18-25). Friedr. Vieweg & Sohn Verlag. https://doi.org/10.1007/978-3-8348-9418-2_2
- Samarati, P., & Sweeney, L. (1998). Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. *Proceedings of the 1998 IEEE Symposium on Research in Security and Privacy, USA*, 1-19. https://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf
- Sax, M. (2018). Privacy from an ethical Perspective. In B. Van der Sloot & A. De Groot (Eds.), *The Handbook of Privacy Studies: An Interdisciplinary Introduction* (pp. 143-173). Amsterdam University Press.
- Schaar, P. (2010). Privacy by design. *Identity in the Information Society*, 3, 267-274. <https://doi.org/10.1007/s12394-010-0055-x>
- Schermer, B. W. (2011). The limits of privacy in automated profiling and data mining. *Computer Law & Security Review*, 27(1), 45-52. <https://doi.org/10.1016/j.clsr.2010.11.009>
- Schneider, G. (2018). Is privacy by construction possible? In T. Margaria & B. Steffen (Eds.), *Leveraging applications of formal methods, verification and validation. Modeling. Lecture Notes in Computer Science (including subseries Theoretical Computer Science and General Issues)* (Vol. 11244, pp. 471-485). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-03418-4_28
- Schneier, B. (2009). Architecture of privacy. *IEEE Security & Privacy*, 7(1), 88-88. <https://doi.org/10.1109/MSP.2009.1>
- Schneier, B. (2015). *Data and Goliath: The hidden battles to capture your data and control your world*. W. W. Norton & Company.
- Schütz, P., & Friedewald, M. (2011). Privacy: What are we actually talking about?. In S. Fischer-Hübner, P. Duquenoy, M. Hansen, R. Leenes, & G. Zhang (Series Eds.), *Privacy and identity management for life. Privacy and Identity 2010. IFIP Advances in Information and Communication Technology* (Vol. 352, pp. 1-14). Springer. https://doi.org/10.1007/978-3-642-20769-3_1
- Schwartz, P. M. (1999). Privacy and democracy in cyberspace. *Vanderbilt Law Review*, 52, 1609-1701. <https://escholarship.org/uc/item/2fq3v1mj>
- Schwarz, N. (1990). Feeling as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognitions: Foundations of social behavior*, Vol. 2 (p. 527-561). The Guilford Press.

- Schwarz, N. (2012). Feelings-as-information theory. In P. van Lange, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (pp. 289-308). Sage Publications Ltd.
- Sedenberg, E., Chuang, J., & Mulligan, D. (2016). Designing commercial therapeutic robots for privacy preserving systems and ethical research practices within the home. *International Journal of Social Robotics*, 8, 575-587. <https://doi.org/10.1007/s12369-016-0362-y>
- Sharp, T. (2013, June 12). *Right to privacy: Constitutional rights & privacy laws*. Live Science. <https://www.livescience.com/37398-right-to-privacy.html>
- Shen, B., Wei, L., Xiang, C., Wu, Y., Shen, M., Zhou, Y., & Jin, X. (2021). Can systems explain permissions better? Understanding users' misperceptions under smartphone runtime permission model. *Proceedings of 30th Usenix Security Symposium*. Advanced online publication. <https://www.usenix.org/conference/usenixsecurity21/presentation/shen-bingyu>
- Shen, Y., & Vervier, P.-A. (2019). IoT security and privacy labels. In M. Naldi, G. F. Italiano, K. Rannenberg, M. Medina, & A. Bourka (Eds.), *Privacy Technologies and Policy. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11498, pp. 136-147). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-21752-5_9
- Sheng, X., Felix, R., Saravade, S., Siguaw, J. A., Ketron, S. C., Kreijtz, K., & Duchowski, T. (2020). Sight unseen: The role of online security indicators in visual attention to online privacy information. *Journal of Business Research*, 111, 218-240. <https://doi.org/10.1016/j.jbusres.2019.11.084>
- Shi, P., Xu, H., & Chen, Y. (2013). Using contextual integrity to examine interpersonal information boundary on Social Network Sites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, France*, 35-38. <https://doi.org/10.1145/2470654.2470660>
- Shklovski, I., Mainwaring, S. D., Skúladóttir, H. H., & Borgthorsson, H. (2014). Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Canada*, 2347-2356. <https://doi.org/10.1145/2556288.2557421>
- Shvartzshnaider, Y., Wies, T., Kift, P., Nissenbaum, H., Subramanian, L., & Mittal, P. (2016). Learning privacy expectations by crowdsourcing contextual informational norms. *Proceedings of the Fourth AAAI Conference on Human Computation and Crowdsourcing, USA*, 4(1), 209-218. <https://ojs.aaai.org/index.php/HCOMP/article/view/13271>
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70, 747-770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- Simitis, S. (1987). Reviewing privacy in an information society. *University of Pennsylvania Law Review*, 135(3), 707-746. <https://doi.org/10.2307/3312079>

- Simmel, G. (1992). Das Geheimnis und die geheime Gesellschaft. In O. Rammstedt (Ed.), *Gesamtausgabe Band 11, Soziologie: Untersuchungen über die Formen der Vergesellschaftung* (pp. 383-455). Suhrkamp.
- Simon, H. A. (1955). A behavioural model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118. <https://doi.org/10.2307/1884852>
- Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.), *Decision and Organization* (pp. 161-176). North-Holland Publishing Company.
- Simon, H. A. (1982). *Models of bounded rationality*. MIT Press.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41, 1-20. <https://doi.org/10.1146/annurev.ps.41.020190.000245>
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.
- Smith, E. J., & Kollars, N. A. (2015). QR panopticism: user behavior triangulation and barcode-scanning applications. *Information Security Journal: A Global Perspective*, 24(4–6), 157-163. <https://doi.org/10.1080/19393555.2015.1085113>
- Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4), 989-1015. <https://doi.org/10.2307/41409970>
- Smith, K. L., Shade, L. R., & Shepherd, T. (2017). Open privacy badges for digital policy literacy. *International Journal of Communication*, 11, 2784-2805.
- Solove, D. J. (2002). Conceptualizing privacy. *California Law Review*, 90, 1087–1155. https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2086&context=faculty_publications
- Solove, D. J. (2006). A brief history of information privacy law. In K. J. Mathews (Ed.), *Proskauer on Privacy: A Guide to Privacy and Data Security Law in the Information Age. GWU Law School Public Law Research, paper No. 215* (pp. 1-53). Practising Law Institute.
- Solove, D. J. (2008). Understanding privacy. *GWU Legal Studies Research Paper No. 420, GWU Law School Public Law Research Paper No. 420*. Harvard University Press. <https://ssrn.com/abstract=1127888>
- Solove, D. J., & Schwartz, P. M. (2011). *Privacy law fundamentals*. GWU Law School Public Law Research Paper No. 542, UC Berkeley Public Law Research Paper No. 1790262, GWU Legal Studies Research Paper No. 542. International Association of Privacy Professionals. <https://ssrn.com/abstract=1790262>
- Solove, D. J., & Schwartz, P. M. (2021). *Consumer privacy and data protection*. Wolters Kluwer.
- Solove, D. J., Rotenberg, M., & Schwartz, P. M. (2006). *Privacy, information, and technology*. Aspen Publishers, Inc.
- Soumelidou, A., & Tsohou, A. (2019). Effects of privacy policy visualization on users' information privacy awareness level: The case of Instagram. *Information Technology & People*, 33(2), 502-534. <https://doi.org/10.1108/ITP-08-2017-0241>
- Spdr870. (2017). Mellow Meadows Tower Defense. Google Play Store. Abandoned.

- Spensky, C., Stewart, J., Yerukhimovich, A., Shay, R., Trachtenberg, A., Housley, R., & Cunningham, R. K. (2016). SoK: Privacy on mobile devices – it's complicated. *Proceedings on Privacy Enhancing Technologies*, 2016(3), 96-116. <https://doi.org/10.1515/popets-2016-0018>
- Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K.-L. (2015). The challenges of personal data markets and privacy. *Electronics Markets*, 25, 161-167. <https://doi.org/10.1007/s12525-015-0191-0>
- Sproull, L., & Kiesler, S. (1986). Reducing social context cues: electronic mail in organizational communication. *Management Science*, 32(11), 1492-1512. <https://doi.org/10.1287/mnsc.32.11.1492>
- Sproull, L., & Kiesler, S. (1991). *Connections: New ways of working in the networked organization*. MIT Press.
- Stanton, J. M. (2003). Information technology and privacy: A boundary management perspective. In S. Clarke, E. Coakes, M. G. Hunter, & A. Wenn (Eds.), *Socio-technical and human cognition elements of information systems* (pp. 79-103). Idea Group Publishing. <https://doi.org/10.4018/978-1-59140-104-9.ch005>
- Statista. (2016, November, 23). *Number of mobile phone users worldwide from 2015 to 2020 (in billions)*. <https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/>
- Statista. (2019, November, 18). *Number of mobile app downloads worldwide from 2016, 2017, and 2021, by store (in billions)*. <https://www.statista.com/statistics/276602/annual-number-of-mobile-app-downloads-by-store/>
- Stephenson, W. (1953). *The study of behaviour: Q-technique and its methodology*. University of Chicago Press.
- Stojkovski, B., & Lenzini, G. (2020). Evaluating ambiguity of privacy indicators in a secure email app. In M. Loreti & L. Spalazzi (Eds.), *Proceedings of the Fourth Italian Conference on Cyber Security, Italy. CEUR Workshop Proceedings* (Vol. 2597, pp. 223-234). CEUR-WS.org. <http://hdl.handle.net/10993/43267>
- Sundar, S. S., Kang, H., Wu, M., Go, E., & Zhang, B. (2013). Unlocking the privacy paradox: Do cognitive heuristics hold the key?. *Proceedings of CHI'13 Extended Abstracts on Human Factors in Computing Systems, France*, 811-816. <https://doi.org/10.1145/2468356.2468501>
- Sunyaev, A., Dehling, T., Taylor, P. L., & Mandl, K. D. (2015). Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*, 22(e1), e28-e33. <https://doi.org/10.1136/amiajnl-2013-002605>
- Suphakul, T., & Senivongse, T. (2017). Development of privacy design patterns based on privacy principles and UML. *Proceedings of the 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Japan*, 369-375. <https://doi.org/10.1109/SNPD.2017.8022748>

- Sutanto, J., Palme, E., Tan, C.-H., & Phang, C. W. (2013). Addressing the personalization-privacy paradox: an empirical assessment from a field experiment on smartphone users. *MIS Quarterly*, 37(4), 1141-1164. <https://doi.org/10.25300/misq/2013/37.4.07>
- Sweeney, L. (2002). *k*-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5), 557-570. <https://doi.org/10.1142/S0218488502001648>
- Taddicken, M. (2014). The 'privacy paradox' in the social Web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, 19(2), 248-273. <https://doi.org/10.1111/jcc4.12052>
- Tamburri, D. A. (2020). Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation. *Information Systems*, 91, article 101469, 1-14. <https://doi.org/10.1016/j.is.2019.101469>
- Tan, F. T. C., & Vasa, R. (2011). Toward a social media usage policy. In P. Seltsikas, D. Bunker, L. Dawson, & M. Indulska, M. (Eds.), *Proceedings of the 22nd Australasian Conference on Information Systems - Identifying the Information Systems Discipline, Australia*, 72, 1-11. AIS. <http://aisel.aisnet.org/acis2011/72>
- Taylor, H. (2003). *Most people are "privacy pragmatists" who, while concerned about privacy, will sometimes trade it off for other benefits. The Harris Poll #17, March 19, 2003*. Harris Interactive. <https://theharrispoll.com/wp-content/uploads/2017/12/Harris-Interactive-Poll-Research-Most-People-Are-Privacy-Pragmatists-Who-While-Conc-2003-03.pdf>
- Taylor, V. F., & Martinovic, I. (2016). *Quantifying permission-creep in the Google Play Store*. *ArXiv*, *abs/1606.01708*. <https://arxiv.org/pdf/1606.01708.pdf>
- Technical Committee ISO/IEC JTC 1SC 27. (2011). ISO/IEC 29100:2011. Information technology – Security techniques – Privacy framework. International Organization for Standardization and International Electrotechnical Commission. <https://www.iso.org/standard/45123.html>
- Tene, O., & Polonetsky, J. (2011). Privacy in the age of big data: a time for big decisions. *Stanford Law Review Online*, 64, 63-69.
- Tesfay, W. B., Hofmann, P., Nakamura, T., Kiyomoto, S., & Serna, J. (2018). PrivacyGuide: Towards an implementation of the EU GDPR on internet privacy policy evaluation. *Proceedings of the Fourth ACM International Workshop on Security and Privacy, USA*, 15-21. <https://doi.org/10.1145/3180445.3180447>
- The Privacy Company B.V. (2019). *Privacy by design framework. Versie 3.0 oktober 2019*. https://uploads-ssl.webflow.com/5d5d0a009052fec16249aaab/5dc3e8a351595bfff2a94fda_Privacy%20by%20Design%20framework%20V3.pdf
- Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, 93, 1-12. <https://doi.org/10.1016/j.chb.2018.11.046>

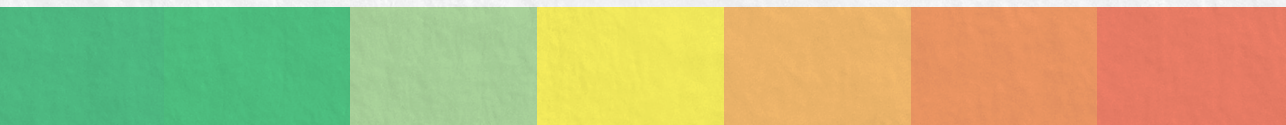
- Tokas, S., Owe, O., & Ramezanifarkhani, T. (2020). Language-based mechanisms for privacy-by-design. In M. Friedewald, M. Önen, E. Lievens, S. Krenn, & S. Fricker (Series Eds.), *Privacy and identity management. Data for better living: AI and privacy. IFIP Advances in Information and Communication Technology* (Vol. 576, pp. 142-158). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-42504-3_10
- Tönnies, F. (2012). *Studien Zu Gemeinschaft Und Gesellschaft*. Springer VS. <https://doi.org/10.1007/978-3-531-94174-5>
- Trepte S., & Masur P. K. (2020). Need for Privacy. In V. Zeigler-Hill, & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences*. Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-319-24612-3_540
- Trepte, S. (2016). The paradoxes of online privacy. In: M. Walrave, K. Ponnet, E. Vanderhoven, J. Haers, & B. Segaert (Eds.), *Youth 2.0: Social Media and Adolescence* (pp. 103-115). Springer. https://doi.org/10.1007/978-3-319-27893-3_6
- Tsai, J. Y., Cranor, L., Acquisti, A., & Fong, C. (2006). What's it for you? A survey of online privacy concerns and risk. *NET Institute Working Paper, No. 06-29*, 1-21. <http://dx.doi.org/10.2139/ssrn.941708>
- Tsai, J. Y., Egelman, S., Cranor, L. F., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research, 22*(2), 254-268. <https://doi.org/10.1287/isre.1090.0260>
- Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In D. Wendt, & C. Vlek (Eds.), *Utility, Probability, and Human Decision Making. Theory and Decision Library (An International Series in the Philosophy and Methodology of the Social and Behavioral Sciences)* (Vol. 11, pp. 141-162). Springer. https://doi.org/10.1007/978-94-010-1834-0_8
- U.S. Const. amend. I.
- U.S. Const. amend. III
- U.S. Const. amend. IV
- U.S. Const. amend. V
- Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019). (Un)informed consent: Studying GDPR consent notices in the field. *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, United Kingdom*, 973-990. <https://doi.org/10.1145/3319535.3354212>
- Van den Berg, B., & Van der Hof, S. (2012). What happens to my data? A novel approach to informing users of data processing practices. *First Monday, 17*(7), 1-15. <https://doi.org/10.5210/fm.v17i7.4010>
- Van Der Sype, Y. S., & Maalej, W. (2014). On lawful disclosure of personal user data: What should app developers do?. *Proceedings of the 7th International Workshop on Requirements Engineering and Law, Karlskrona*, 25-34. <https://doi.org/10.1109/RELAW.2014.6893479>

- Van Kleek, M., Liccardi, I., Binns, R., Zhao, J., Weitzner, D. J., & Shadbolt, N. (2017). Better the devil you know: Exposing the data sharing practices of smartphone apps. *Proceedings of the CHI Conference on Human Factors in Computing Systems, USA*, 5208-5220. <https://doi.org/10.1145/3025453.3025556>
- Van Kranenburg, R. (2011). The Internet of Things. *World Affairs: The Journal of International Issues*, 15(4), 126-141. <https://www.jstor.org/stable/48505085>
- Vanezi, E., Kouzapas, D., Kapitsaki, G. M., Costi, T., Yeratziotis, A., Mettouris, C., Philippou, A., & Papadopoulos, G. A. (2019). GDPR compliance in the design of the INFORM e-Learning platform: A case study. In M. Kolp, J. Vanderdonckt, M. Snoeck, & Y. Wautelet (Eds.), *Proceedings of the 13th International Conference on Research Challenges in Information Science, Belgium* (pp. 1-12). IEEE. <https://doi.org/10.1109/RCIS.2019.8877022>
- Varian, H. R. (2002). Economic aspects of personal privacy. In W. H. Lehr & L. M. Pupillo (Series Eds.), *Cyber Policy and Economics in an Internet Age. Topics in regulatory economics and policy series* (Vol. 43, pp. 127-137). Springer. https://doi.org/10.1007/978-1-4757-3575-8_9
- Vasa, R., Hoon, L., Mouzakis, K., & Noguchi, A. (2012). A preliminary analysis of mobile app user reviews. In V. Farrell, G. Farrell, C. Chua, W. Huang, R. Vasa, & C. Woodward (Eds.), *Proceedings of the 24th Australian Computer-Human Interaction Conference, Australia* (pp. 241-244). ACM. <https://doi.org/10.1145/2414536.2414577>
- Veltri, G. A., & Ivchenko, A. (2017). The impact of different forms of cognitive scarcity on online privacy disclosure. *Computers in Human Behavior*, 73, 238-246. <https://doi.org/10.1016/j.chb.2017.03.018>
- Vemou, K., & Karyda, M. (2014). Guidelines and tools for incorporating privacy in social networking platforms. *IADIS International Journal on WWW/Internet*, 12(2), 16-33. <http://www.iadisportal.org/ijwi/papers/2014121202.pdf>
- Vickery, A. (2008). An Englishman's home is his castle? Thresholds, boundaries and privacies in the eighteenth-century London house. *Past & Present*, 199(1), 147-173, <https://doi.org/10.1093/pastj/gtn006>
- Vincent, D. (2016). *Privacy: a short history*. Polity Press.
- Volkamer M., Renaud K., Kulyk O., & Emeröz S. (2015). A socio-technical investigation into smartphone security. In S. Foresti (Ed.), *Security and trust management. Lecture Notes in Computer Science* (Vol. 9331, pp. 265-273). Springer. https://doi.org/10.1007/978-3-319-24858-5_17
- Vroom, V. H. (1964). *Work and Motivation*. Wiley.
- Wagner, I., & Eckhoff, D. (2018). Technical privacy metrics: A systematic survey. *ACM Computing Surveys*, 51(3), article 57, 1-38. <https://doi.org/10.1145/3168389>
- Wakefield, R. L. (2013). The influence of user affect in online information disclosure. *The Journal of Strategic Information Systems*, 22(2), 157-174. <https://doi.org/10.1016/j.jsis.2013.01.003>

- Wakefield, R. L., & Whitten, D. (2006). Mobile computing: a user study on hedonic/ utilitarian mobile device usage. *European Journal of Information Systems*, 15, 292-300. <https://doi.org/10.1057/palgrave.ejis.3000619>
- Waldman, A. E. (2020). Cognitive biases, dark patterns, and the privacy paradox. *Current Opinion in Psychology*, 31, 105-109. <https://doi.org/10.1016/j.copsyc.2019.08.025>
- Wang, H., Lee, M. K. O., & Wang, C. (1998). Consumer privacy concerns about Internet marketing. *Communications of the ACM*, 41(3), 63-70. <https://doi.org/10.1145/272287.272299>
- Wang, T., Duong, T. D., & Chen, C. C. (2016). Intention to disclose personal information via mobile applications: a privacy calculus perspective. *International Journal of Information Management*, 36(4), 531-542. <https://doi.org/10.1016/j.ijinfomgt.2016.03.003>
- Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review*, 4(5), 193-220. <https://doi.org/10.2307/1321160>
- Westin, A. F. (1967). *Privacy and freedom*. Atheneum.
- Westin, A. F. (2003). Social and political dimensions of privacy. *Journal of Social Issues*, 59(2), 431-453. <https://doi.org/10.1111/1540-4560.00072>
- Williams, M., Nurse, J. R. C., & Creese, S. (2017). Privacy is the boring bit: user perceptions and behaviour in the Internet-of-Things. *Proceedings of the 15th International Conference on Privacy, Security and Trust, Canada*, 181-190. <https://doi.org/10.1109/PST.2017.00029>
- Wilson, D. W., & Valacich, J. S. (2012). Unpacking the privacy paradox: Irrational decision-making within the privacy calculus. *Proceedings of the Thirty Third International Conference on Information Systems, USA*, 1-11. <https://aisel.aisnet.org/icis2012/proceedings/ResearchInProgress/101/>
- Wondracek, G., Holz, T., Kirda, E., & Kruegel, C. (2010). A practical attack to de-anonymize social network users. *Proceedings of the 2010 IEEE Symposium on Security and Privacy, USA*, 223-238. <https://doi.org/10.1109/SP.2010.21>
- Xu, H., Dinev, T., Smith, H. J., & Hart, P. (2008). Examining the formation of individual's privacy concerns: Toward an integrative view. *Proceedings of the Twentey Ninth International Conference on Information Systems, France*, paper 6, 1-16. <https://aisel.aisnet.org/icis2008/6>
- Xu, H., Gupta, S., Rosson, M. B., & Carroll, J. M. (2012). Measuring mobile users' concerns for information privacy. *Proceedings of the Thirty Third International Conference on Information Systems, USA*, 1-16. <https://aisel.aisnet.org/icis2012/proceedings/ISSecurity/10/>
- Xu, H., Teo, H.-H., Tan, B. C. Y., & Agarwal, R. (2010). The role of push-pull technology in privacy calculus: the case of location-based services. *Journal of Management Information Systems*, 26(3), 135-173. <https://doi.org/10.2753/MIS0742-1222260305>
- Yang, H. C. (2013). Bon appétit for apps: Young American consumers' acceptance of mobile applications. *Journal of Computer Information Systems*, 53(3), 85-95. <https://doi.org/10.1080/08874417.2013.11645635>

- Yoo, C. W., Ahn, H. J., & Rao, H. R. (2012). An exploration of the impact of information privacy invasion. *Proceeding of Thirty Third International Conference on Information Systems, USA*, 1-18. <https://aisel.aisnet.org/icis2012/proceedings/ISSecurity/2/>
- Young, A. L., & Quan-Haase, A. (2013). Privacy protection strategies on Facebook. *Information, Communication & Society*, 16(4), 479-500. <https://doi.org/10.1080/1369118X.2013.777757>
- Young, S. (2021). Zoombombing your toddler: User experience and the communication of Zoom's privacy crisis. *Journal of Business and Technical Communication*, 35(1), 147-153. <https://doi.org/10.1177/1050651920959201>
- Yu, L., Zhang, T., Luo, X., & Xue, L. (2015). Autoppg: Towards automatic generation of privacy policy for android applications. *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices, USA*, 39-50. <https://doi.org/10.1145/2808117.2808125>
- Zafeiropoulou, A. M., Millard, D. E., Webber, C., & O'Hara, K. (2013). Unpicking the privacy paradox: Can structuration theory help to explain location-based privacy decisions?. *Proceedings of the 5th Annual ACM Web Science Conference, France*, 463-472. <https://doi.org/10.1145/2464464.2464503>
- Zang, J., Dummit, K., Graves, J., Lisker, P., & Sweeney, L. (2015). Who knows what about me? A survey of behind the scenes personal data sharing to third parties by mobile apps. *Technology Science, 2015103001*. <http://techscience.org/a/2015103001>
- Zarsky, T. Z. (2017). Incompatible: The GDPR in the age of big data. *Seton Hall Law Review*, 47, No. 4(2), 995-1018. <https://ssrn.com/abstract=3022646>
- Zarsky, T. Z. (2019). Privacy and manipulation in the digital age. *Theoretical Inquiries in Law*, 20(1), 157-188. <https://doi.org/10.1515/til-2019-0006>
- Zhang, B., Wu, M., Kang, H., Go, E., & Sundar, S. S. (2014). Effects of security warnings and instant gratification cues on attitudes toward mobile websites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Canada*, 111-114. <https://doi.org/10.1145/2556288.2557347>
- Zhauniarovich, Y., & Gadyatskaya, O. (2016). Small changes, big changes: An updated view on the Android permission system. In F. Monrose, M. Dacier, G. Blanc, & J. Garcia-Alfaro (Series Eds.), *Research in Attacks, Intrusions, and Defenses. Lecture Notes in Computer Science* (Vol. 9854, pp. 346-367). Springer. https://doi.org/10.1007/978-3-319-45719-2_16
- Zimmerman, S., Thorpe, A., Fox, C., & Kruschwitz, U. (2019). Investigating the interplay between searchers' privacy concerns and their search behavior. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, France*, 953-956. <https://doi.org/10.1145/3331184.3331280>
- Zorzo, S., Pontes, D., Mello, J., & Dias, D. (2016). Privacy rules: Approach in the label or textual format. *Proceedings of the 22nd Americas Conference on Information Systems: Surfing the IT Innovation Wave, USA*, 1-10. <https://aisel.aisnet.org/amcis2016/ISSec/Presentations/44/>

Zukowski, T., & Brown, I. (2007). Examining the influence of demographic factors on internet users' information privacy concerns. *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries, South Africa*, 197–204. <https://doi.org/10.1145/1292491.1292514>



Appendices

Appendix 2.1 *Definition of Theories (in alphabetical order)*

Theory	Definition
Adaptive Cognition Theory of Social Network Participation (Hu & Ma, 2010)	Users' participation in SNs consists of three phases: initial use, exploratory use and managed use. The progression from one phase to the next results from the understanding of benefits and risks and the adaptation of activities and controls. The final phase is relatively stable but can easily be damaged or altered by negative experiences or positive reinforcements. Based on rational choice theory and behaviorism.
Cognitive Heuristics (Tversky & Kahneman, 1975)	Rule-of-thumb strategies play an important role in decision-making. These mental shortcuts allow individuals to come to a decision quickly without the urge to think about the next action. But heuristics can lead to biases as something that has been applicable in the past, is not necessarily suitable in another situation. Furthermore, heuristic hinders individuals from developing new ideas and alternative solutions.
Communication Privacy Management Theory (Petronio, 1991, 2002)	The decision on which information to reveal and which to keep private. Publication or retention goes along with certain risk and benefits. This process is guided by the subjective privacy boundaries individuals have (determined by an iterative process of rule development, boundary coordination, boundary turbulence) and reshaped continuously, depending on situation, context and communication partner(s).
Conformity and Peer Group Pressure (Crutchfield, 1955)	Individuals feel indirect pressure to change their own behavior in order to conform to an admired peer group. Peer group pressure can either result in positive or negative reactions (e.g., start smoking or studying regularly).
Cues-filtered-out Theory (Sproull & Kiesler, 1986, 1991)	This theory implies that individuals disclose more personal data in computer-mediated communication settings compared to face-to-face settings due to the absence of social and contextual cues.

Appendix 2.1 *Continued.*

Theory	Definition
Duality of Gemeinschaft und Gesellschaft (Tönnies, 2012)	Some forms of social collectives are determined by internalized emotional ties and implicit rules (Gemeinschaft), whereas other collectives are determined by rational calculations and explicit rules (Gesellschaft). In social networks, people share private information because this is an implicit rule for belonging to a certain group. Although people know about data violation (explicitly) albeit on a very abstract level, these rational feelings cannot be translated into actual feelings of fear (Gesellschaft). As a result, the feeling of belonging to a social network overpowers the threats of data misuse.
Dual Process Model of Cognition (Kahneman, 2003)	Decision-making is based on two systems. System I is fast and automatic but is vulnerable to influences that inhibit the rational decision-making process (produces intuitive concern for instance), whereas System II is rational and responsible for reasoning (produces considered concern, for instance). According to this theory, there are two explanations for the privacy paradox: Either individuals act on their intuitive concern without assessing the risks due to incomplete understanding of it (no considered concern takes place) or high considered concern may be overridden by low considered concern (privacy concerns are considered but individuals are unable to address them adequately).
Expectancy Theory (Vroom, 1964)	Behavior is a result of conscious choices with the purpose to maximize gain and minimize loss. This decision-making process is based on three beliefs: valence (emotional attitude toward outcome and strength of wanting a reward), expectancy (self-confidence to do s.th.) and instrumentality (perception of probability for gaining reward). Based on these beliefs, an individual chooses a certain behavior over others because of the expected outcome of that specific behavior.

Appendix 2.1 *Continued.*

Theory	Definition
Extended Two-Component Model of Self-Presentation Online (based on Leary & Kowalski, 1990)	Self-presentation (impression management) is the process by which individuals try to control the impression they make on others. This process is determined by two components: impression motivation (the willingness to create or re-create impressions in another's mind; influenced by goal relevance, value of desired goal and discrepancy between desired and current self-image) and impression construction: the process of creating this impression through change of behavior; influenced by self-concept, desired identity, role constraints, current or potential social image and target values.
Extension to the Privacy Calculus Theory (Culnan & Armstrong, 1999)	The privacy paradox may result from misleading situational cues which bias the cognitive valuation processes (e.g., affective thinking) and the prevalence of situation-specific considerations as compared to generic attitudes. Eventually, privacy disclosure intention is determined by situational cues even if dispositional attitudes regarding privacy behavior are different to intention. The study of Kehr et al. (2015) showed that privacy decisions are driven by situation-specific privacy assessment, general dispositions (i.e. general privacy concerns, institutional trust) and affect-based heuristics (quite often subconscious processes).
Feelings-as-Information Theory (Schwarz, 1990, 2012)	Individuals rely on their feelings (mood, meta-cognition, emotion and body sensation) in decision-making processes, often leading to accurate responses but sometimes not. While in a good mood for example, people evaluate targets or situations as more positively. Furthermore, judgments are based on feelings of ease or difficulty as situations or targets which are easy to process are evaluated more positively, less risky and more valuable.
Hyperbolic Discounting Theory (Laibson, 1997)	If there is a choice, individuals usually choose a small benefit in the short term over a larger benefit in the longer term. If all choices are available on the long term, larger benefits will be chosen, even if these will occur later than smaller benefits.
Immediate Gratifications (O'Donoghue & Rabin, 2001)	Quite often, individuals encounter self-control problems due to immediate gratifications (present bias) which leads to behavior that may backfire in the long run.

Appendix 2.1 *Continued.*

Theory	Definition
Optimistic Bias Theory (Irwin, 1953)	Individuals have a tendency to underestimate the likelihood of experiencing adverse events. This might result in denying precautions which might lower risk perception.
Privacy Calculus Theory (Culnan & Armstrong, 1999)	The intention to disclose personal information is based on a rational risk-benefit calculation as perceived benefits are weighed against risk probability. If perceived benefits outweigh risks, information might be disclosed in exchange for social or economic benefit.
Privacy Regulation Theory (Altman, 1975)	Privacy is a dynamic process of interaction regulation with others. Based on internal states and external conditions, individuals determine the degree of openness. Privacy regulation should be done at an optimal level (desired level of privacy is equal to actual level). Here, trust plays an important role in the interaction regulation process which is defined by self-boundary (around a person) which is modified by self-disclosure and a dyadic boundary (ensures discloser's safety in case of violation).
Prospect Theory (Kahneman & Tversky, 1979)	Individuals do not process information in a rational way. Decision-making processes take place in two stages. During the editing stage, expected outcomes are ordered based on heuristics by setting a reference point. During the evaluation stage, outcomes lesser than the reference point are considered as losses and greater outcomes as gains. Furthermore, losses are more heavily weighted than an equal amount of gains.
Public Value Theory (Meynhardt, 2009)	In Public Value Theory, any organization contributes to a society's wellbeing (objective facts) as long as such individuals perceive their relationship to the public either positively or negatively (objective facts are reflected in people's perceptions and subjective evaluation). If an organization is perceived as trustworthy regarding data protection but their public value is low, this organization does not contribute to the public value, unless data protection is valued by the public. This explains partly the privacy paradox as people do not engage in protective behavior because they fail to value data protection adequately.

Appendix 2.1 Continued.

Theory	Definition
Quantum Theory (Based on Busemeyer, Wang, & Townsend 2006)	Objective reality does not exist. An object can take up any possible states simultaneously as long as an individual does not evaluate it. The evaluation changes the object's state.
Rational Choice Theory of Human Behavior (Simon, 1955)	Decisions are always reasonable and logical in order to gain the greatest benefit or satisfaction in an individual's self-interest.
Rational Ignorance Theory (Downs, 1957)	The conscious choice of an individual to not pay attention to certain information is based on a cost-benefit calculation (costs of learning are disproportionate to potential benefits).
Resource Exchange Theory (Donnenwerth & Foa 1974; Foa, 1971)	Individuals try to rationally exchange resources with others due to their wishes and needs. Furthermore, through participation in a social system, individuals may contribute to a certain group and get benefits from each other. In exchange for other resources such as money, services, time, status and love (e.g., online relationships), people are willing to provide personal resources (e.g., personal information).
Self-Control Bias (Loewenstein, 1999)	The tendency to favour immediate rewards on the short term at the expense of future risks due to lack of self-discipline.
Symbolic Interactionism (Blumer, 1986)	Social interaction creates and maintains social structures and meanings. By interacting with others over time, people share meaning and actions and come to understand events in certain and similar ways. This is the basis for society.
Social Representation Perspective (Abric, 1996; Moscovici, 1984)	Social representations are values, ideas or practices that enable individuals to orient and master the social world. By means of social exchange, new concepts are integrated into existing representations (making the unfamiliar familiar) by means of anchoring (the fit of new knowledge into existing representation is proven through anchoring) and objectification (make abstract concepts concrete via the creation of a new representation, e.g., the concept of privacy).

Appendix 2.1 *Continued.*

Theory	Definition
Structuration Theory (Giddens, 1984)	Social life is more than individual acts, yet it is determined by social forces such as traditions, institutions or moral codes. Social structures determine human behavior but can also be altered due to perception differences, ignorance or replacement. Therefore, behavior is a balance between social structures and agency, known as the ability to act on one's own free will. The structure is achieved through a dynamic process: Structure forms the basis for decision-making but is at the same time the outcome of it.
Theory of Bounded Rationality (Simon, 1982)	Quite often, individuals are satisfied with a solution that is good enough but not optimal due to cognitive limitations as they are unable to access and process all of the information that would be needed to do so. Even with all of the information at hand, cognitive processing would be impossible. Individuals constantly try to rationally maximize benefits but decision-making can only be rational within the limits of cognitive ability and available time.
Theory of Cognitive Absorption (Agarwal & Karahanna, 2000; Agarwal et al. 1997)	Individuals in the flow state, called cognitive absorption, suppress processes that call someone's attention to feelings of cognitive dissonance, leading to inappropriate privacy calculation.
Theory of Incomplete Information (Harsanyi, 1967)	In game theory, one party is less informed than the other or in other words, not all parties know each other's utilities and rules.
Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980)/Theory of Planned Behavior (TPB) (Ajzen, 1985)	According to TRA, an individual's behavioral intention depends on their attitude toward a certain behavior and the subjective norms. The stronger the intention, the more likely that a person will engage in a certain behavior. The TPB also predicts the likelihood of an individual's intention to engage in a certain self-controlled behavior. Behavioral intention is influenced by existing attitudes as they pertain to desired outcomes, social norms and the evaluation of the risk-to-benefit ratio of that outcome (perceived behavioral control). The stronger the attitudes, perceived control and compliance with social norms, the more likely it is that individuals will engage in certain behaviors.

Appendix 2.1 *Continued.*

Theory	Definition
Theory of Ritualized Media Use (Rubin, 1984)	The use of media extends beyond satisfying information and entertainment needs to be seen as a habitual pastime that is integrated into everyday life routines that are connected to temporary structures (favorite show at a particular time) and social rituals (meeting friends to watch a show).
Theory of Under Insurance (Kunreuther, 1984)	The tendency toward reluctance to engage in privacy protection behavior against low probability but high impact/consequences events due to biased perception (event is less threatening than it actually is), underestimation of probability (as a consequence of little or no experience with the threat in question), unawareness of the threat or costs for engagement are considered as too high.
Third-Person Effect Theory (Davison, 1983)	Individuals tend to overestimate the effect of media on others while underestimating the influence on themselves (due to social desirability = denying influence goes along with self-esteem, creating social distance to a certain group and influence is self-chosen by others). As a result, individuals usually do not demonstrate the intended behavior as a response to the message.
Uses and Gratification Theory (Blumler & Katz, 1974; Katz et al., 1974)	Media use is actively determined in order to achieve and satisfy certain goals and needs among the dimension of diversion and entertainment, building and maintaining relationships and identity construction. This assumes that individuals know their needs and how to gratify them. Individuals consume certain media either for process gratification or content gratification.

Appendix 3.1 Selection of apps used for the experimental part of the study

	App	Amount of permissions	Permissions	Degree of intrusiveness
Utilitarian app (to-do-list)	My ToDo List	0	-	not intrusive
	ToDo list – Private Tasks	1	Photos/Media/Files	slightly intrusive
	Tasks+ To Do List Manager	2	Identity; Photos/Media/Files	somewhat intrusive
	List & Notes	4	In-app purchases; Identity; Location; Photo/Media/Files	intrusive
	EveryDay ToDo List Task List	5	Identity; contacts, location, photos/media/files; camera	very intrusive
Hedonic app (game)	Mellow Meadows Tower Defense	0	-	not intrusive
	New Eskimo Defense	2	Photos/media/files; device ID & call information	slightly intrusive
	Astroid Defense Classic	4	Location; photo/media/files; WIFI connection information; Device ID & call information	somewhat intrusive
	Safe the cave: Tower Defense	6	In-app purchases; photos/media/files; microphone; camera; WIFI connection information; Device ID & call information	intrusive
	Tower Defense: Infinite War	7	In-app purchases; identity; contacts; phone; photos/media/file; WIFI connection information; Device ID & call information	very intrusive

Appendix 3.2 Questionnaires

Background information (Questionnaire 1)

1. What is your age?
2. What is your gender? (1) *male* (2) *female*
3. Which study program do you follow?
4. How long have you owned a smartphone:
5. On a daily basis, the average time I spent using mobile apps is:
6. On average, the number of mobile apps I use on a weekly basis is:
7. Which operating system do you use on your mobile phone? (1) *Android* (2) *IOS* (3) *other*
8. Where do you usually look for applications?
9. Was your mobile phone ever lost or stolen? (1) *never* (2) *once* (3) *twice* (4) *more than twice*
10. Do you sometimes lend your mobile phone to others? (1) *yes* (2) *yes but only for a while and if I am present* (3) *never*
11. How many apps have you ever installed **yourself** on your mobile phone? (such as Facebook app, games, ringtones, GPS etc.). Please give an indication: (from Yang, 2013)
12. Which categories of apps do you use? (from Google App Store)

Technical knowledge (Questionnaire 1)

from Androulidakis & Kandus (2011) and Kraus et al. (2014)

1. Do you know where you can find your mobile phone's IMEI (International Mobile station Equipment Identity)? (1) *yes* (2) *no* (3) *I don't know what it is*
2. After (re)starting your mobile phone, do you have to enter a PIN for unlocking your SIM card? (1) *yes* (2) *no* (3) *I don't know*
3. Do you use a PIN code or password to unlock your screen-saver? (1) *yes* (2) *no* (3) *doesn't have such feature* (4) *I don't know*
4. Do you use the Bluetooth function? (1) *yes, switched on and visible* (2) *yes, switched on and invisible* (3) *yes, but only for a specific purpose* (4) *no, switched off* (5) *I don't know*
5. Do you run a antivirus app on your mobile phone? (1) *yes* (2) *no* (3) *I don't know*

6. Do you run a static analysis app on your mobile phone, for instance to monitor malicious code patterns, to inspect control flow between apps, or to review requested permissions?
(1) yes (2) no (3) I don't know
7. Do you store passwords in your mobile phone (e.g., credit card password, ATM password)?
(1) yes, encrypted (2) yes, without encryption (3) no (4) I don't know
8. Do you create backup copies of your phone's data?
(1) yes (2) no (3) I don't know
If yes:
How often do you create backup copies of your phone's data?
open question
9. Do you store sensitive personal data in your mobile phone (e.g., photos, videos, audio recordings)
(1) yes (2) no (3) I don't know

Scale 1-7: *(1) completely disagree to (7) completely agree*

10. Communication through mobile phones is safe.
11. I am aware about how the technical characteristics of my mobile phone affect its security.
12. I know how to protect myself against data misuse while surfing in a public network.
13. I know how my mobile phone can be protected from malicious apps.

Download considerations (Questionnaire 1 and 2)

1. Which aspects do you consider when searching for an app? (1) completely disagree to (7) completely agree

(1) trustworthiness of the app (2) prior experience with the app (3) ratings (4) reviews (5) amount of downloads (6) privacy conditions (e.g., information disclosure) (7) security conditions (e.g., data protection) (8) amount of permissions requested (9) clarity of permissions requested (10) readability of permissions requested (11) if permissions are related to functionality (12) usefulness of the app (13) functionality of the app (14) design of the app (15) recommendation (e.g., from your social group) (16) price of the app (17) familiarity with the app (18) other:

Privacy awareness I: General privacy sensitivity (Questionnaire 1)

Westin Privacy Index - see Kumaraguru & Cranor (2005)

Scale 1-4: (1) strongly disagree to (4) strongly agree

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

Downloading an app

Experimental part

Process of choosing, downloading and installing the app (Questionnaire 2)

1. Which app have you downloaded and are you planning to use for the review?
2. Please describe as extensive as possible the decision-making process you followed while (1) choosing for an app, (2) downloading this app and (3) eventually installing this app.
3. Why have you chosen for this app and not for another app?
4. Which aspects of the app did determine your choice for downloading and installing this app? See point 'download considerations'

Writing an app review

1. Which app have you downloaded and will you use for the review?
2. Where did you get the app from? (1) Apple App Store (2) Android Play Store (3) Amazon Marketplace (4) Other
3. Did you pay for the app? (1) yes (2) no
If yes: How much did you pay the app? (1) 1,55 Euro (2) 0,96 Euro (3) 2,99 Euro (4) 1,46 Euro
4. How often did you use the app during the last days? Please give an indication of the total length of time *in hours*:
5. How satisfied are you with the app? Please give an indication on a scale going from 1 to 10 (1 = not satisfied at all to 10 = completely satisfied)
6. Please write a comprehensive review about the app you have downloaded and used during the last week. Please feel free to consider all factors

you want to discuss about the app (e.g., in terms of usability, design, functionality etc.) *Open question*

Privacy awareness II: Information privacy concerns (Questionnaire 3)

MUIPC, from Xu et al. (2012)

Scale 1-7: (1) *completely disagree* to (7) *completely agree*

Perceived surveillance

1. I am concerned that mobile apps are collecting too much information about me.
2. I am concerned that mobile apps may monitor my activities on my mobile device.

Perceived intrusion

1. I feel that as a result of my using mobile apps, others know about me more than I am comfortable with.
2. I believe that as a result of my using mobile apps, information about me that I consider private is now more readily available to others than I would want.
3. I feel that as a result of my using mobile apps, information about me is out there that, if used, will invade my privacy.

Secondary use of personal information

1. I am concerned that mobile apps may use my personal information for other purposes without notifying me or getting my authorization.
2. When I give personal information to use mobile apps, I am concerned that apps may use my information for other purposes.
3. I am concerned that mobile apps may share my personal information with other entities without getting my authorization.

Appendix 5.1 Scenarios corresponding with health and news context and instruction to participants

Imagine you recently downloaded and installed a [*health/news*] application on your mobile phone.

[*Health app*] This app is a pedometer app that registers the steps you walked throughout the day together with the calories burned, covered distance, elapsed time and pace. For doing this, the app asks you to fill in some personal information about you and to carry your smartphone attached to your body. After the start of the app, you can directly access graphics about your movement behavior and other health related information.

[*News app*] This app offers an overview about local, national and international news articles from different information sources. This overview is adapted to news preferences such as sports, politics, economics, technologies or weather. This apps collects some personal information about you, as well as information about the news articles such as actuality of the news article. After starting the app you will get an overview about the most recent news adapted to your interests.

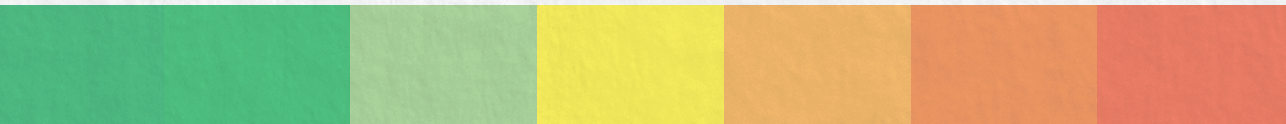
Appendix 5.2 Overview of Q-sets for the contextual factors and corresponding statements

Contextual factor	Statements
<i>What</i>	<p><i>Disclosure of...</i></p> <ul style="list-style-type: none"> • personal pictures • location • place of birth • first name • last name • history of online purchases • contact list • phone number • age • gender • weight • height • number of clicks within app • history of search results • home address • text messages <p>.....</p>

Appendix 5.2 Continued.

Contextual factor	Statements
How	<p><i>Personal data are...</i></p> <ul style="list-style-type: none"> • sold to 3rd parties • used for other purposes than declared • reused to target ads of the app provider • reused to target ads of third parties • transmitted with personal identifiers • stored for a limited time • stored indefinitely • stored within the app environment • stored on an external server • stored on the local device • aggregated from different sources • analyzed for commercial purposes • transmitted without encryption • transmitted with encryption • transmitted with permission of the user • transmitted without permission of the user
Who	<p><i>Personal data are gathered by...</i></p> <ul style="list-style-type: none"> • the SD who programmed the application • the MPP who built the device (e.g., Apple, Samsung, LG) • the provider of the app service • the OS (e.g., iOS by Apple, Android by Google) <p><i>Personal data are analyzed by...</i></p> <ul style="list-style-type: none"> • the SD who programmed the application • the MPP who built the device (e.g., Apple, Samsung, LG) • the provider of the app service • the OS (e.g., iOS by Apple, Android by Google) <ul style="list-style-type: none"> • The SD, who programmed the application... • The OS (e.g., iOS by Apple, Android by Google)... • The MPP, who built the device (e.g., Apple, Samsung, LG)... • The provider of the app service... <p style="padding-left: 40px;"><i>...disseminates my personal data to 3rd parties</i></p> <ul style="list-style-type: none"> • The SD, who programmed the application... • The OS (e.g., iOS by Apple, Android by Google)... • The MPP, who built the device (e.g., Apple, Samsung, LG)... • The provider of the app service... <p style="padding-left: 40px;"><i>...access my personal data</i></p>

Note. SD = software developer; MPP = mobile phone provider; OS = operating system



Summary (in Dutch)

DATA, DATA, AND EVEN MORE DATA:

Empowering users to make well-informed decisions about online privacy

Het internet brengt veel mogelijkheden met zich mee, waaronder connectiviteit en een vrijwel onbeperkte toegang tot informatie. Steeds meer activiteiten in het dagelijks leven verhuizen geheel of gedeeltelijk van offline naar online omgevingen. De toename aan online activiteiten heeft ertoe geleid dat technologieën en achterliggende systemen complexer worden, met nieuwe vragen en dilemma's voor de gebruiker: Welke cookies moet ik accepteren? Zijn willekeurige of gepersonaliseerde advertenties beter? Is een veilig aangemaakt password voldoende of is het beter om twee-factoren authenticatie te gebruiken? De lijst van vragen is eindeloos. Als gebruikers een reële keuze hebben, kiezen ze waarschijnlijk de optie die hun privacy het best beschermt. En dit is de crux van het verhaal: begrijpen gebruikers privacy-instellingen en -documenten? Hebben ze de kennis, vaardigheden en bereidheid om serieus met hun online privacy om te gaan? En in hoeverre zijn ze zich bewust van de consequenties van hun online activiteiten? Bij het gebruik van online services komen gebruikers vaak in situaties waarin ze persoonlijke informatie weggeven hoewel ze bezorgd zijn over hun privacy (Arkko, 2020; Bräunlich et al., 2020; Ochs & Löw, 2012). De discrepantie tussen attitude en gedrag wordt de 'privacy paradox' genoemd. Er bestaan uiteenlopende verklaringen van dit fenomeen. Deze variëren van rationale kosten-baten afwegingen tot non- en irrationele verklaringen (Acquisti, 2004; Deuker, 2010; Keith et al., 2013; Y. Li, 2012). Er is geen consensus over de manier waarop de kloof tussen privacyvoorkeuren en het daadwerkelijke prijsgeven van informatie kan worden verkleind.

Het hoofddoel van dit proefschrift is een op wetenschappelijk onderzoek gebaseerd concept te ontwikkelen dat gebruikers in staat stelt om weloverwogen beslissingen te nemen over hun online privacy. Dit kan door gebruikers te attenderen op de risico's die kunnen ontstaan door het delen van informatie en te helpen om de privacy-instellingen van online services te begrijpen. Om dit hoofddoel te bereiken zijn er twee doelstellingen geformuleerd.

Het eerste doel van dit proefschrift is het verkrijgen van kennis over het online-privacygedrag van gebruikers en over factoren die het prijsgeven van informatie beïnvloeden. Daartoe zijn er een literatuurstudie en drie empirische studies uitgevoerd. Bij het tweede doel van dit proefschrift staat de ontwerpbenadering centraal, met als centrale vraag hoe de meest relevante privacy-aspecten van online services het best op een visuele manier naar gebruikers kunnen worden gecommuniceerd. Om deze vraag te kunnen beantwoorden is er een tweede

literatuurstudie verricht. De resultaten daarvan zijn vervolgens als input genomen voor het ontwerp van een privacyvisualisatie.

Hoofdstuk 2 beschrijft een systematische review van de literatuur over de privacy paradox. Dit onderzoek is uitgevoerd om theoretische benaderingen te identificeren die de discrepantie tussen attitude en gedrag kunnen verklaren. Gebaseerd op een selectie van 32 wetenschappelijke artikelen, die in totaal 35 theorieën bespreken, is een overkoepelend theoretisch kader ontwikkeld. Hiervoor zijn de theorieën ingedeeld in drie verschillende routes van besluitvorming over het prijsgeven van informatie in een online omgeving: (1) een rationele kosten-baten afweging, waarbij de baten uiteindelijk zwaarder wegen dan de kosten, (2) een kosten-baten afweging die vertekend is door irrationele factoren, en (3) een proces waarbij beslissingen worden genomen zonder of op grond van heel beperkte risicoafwegingen. De meeste theorieën gaan uit van de tweede of derde route. Het blijkt dus dat de privacy paradox veel meer behelst dan een rationale overweging (Acquisti, 2004; Debatin et al., 2009; Deuker, 2010; Dinev, 2014; Flender & Müller, 2012; Gambino et al., 2016; Oetzel & Gonia, 2011). Het is waarschijnlijk dat de afweging van kosten en baten beperkt wordt door cognitieve vertekeningen en heuristieken of dat er zelfs helemaal geen of een zeer beperkte risicoafweging plaatsvindt.

Om verder te gaan dan theoretische verklaringen voor de privacy paradox beschrijft **hoofdstuk 3** een experimenteel onderzoek. Eerder onderzoek suggereert dat het ontbreken van technische kennis, beperkte aandacht voor online privacy en financiële overwegingen kunnen leiden tot het prijsgeven van informatie ondanks zorgen om privacy (Liccardi et al., 2014). Voortbordurend op deze aannames richt dit onderzoek zich op de invloed van technische expertise, privacybewustheid en financiële overwegingen bij het delen van informatie. De deelnemers aan dit onderzoek—technisch opgeleide studenten met een bovengemiddelde kennis van online privacy en een relatief hoge privacybewustheid ($N = 66$)—kregen geld dat ze konden gebruiken voor het aankopen van een mobiele applicatie ('app'). Zowel voor als na het downloadproces werd hun gevraagd naar de factoren die volgens hen tijdens het installatieproces een rol spelen. Hoewel de deelnemers aangaven rekening te houden met de betrouwbaarheid van een app en de permissies die een app vraagt, bepalen de aankoopkosten, de beoordelingen van anderen en het design van de app uiteindelijk de keuze voor een app. Veel van de deelnemers eindigden dan ook met een app die niet goed scoorde op privacy-aspecten. Concluderend kan gezegd worden dat technische kennis en een algemene privacybewustheid gebruikers niet beschermen voor ondoordachte beslissingen.

Hoofdstuk 4 beschrijft een vervolgonderzoek naar de relatie tussen technische kennis en aandacht voor online privacy. Hiervoor zijn 20 interviews gehouden met privacy- en cyberbeveiligingsexperts. De interviews gingen over de manier waarop de experts zelf omgaan met hun online. Op grond van hun opvattingen over privacy kunnen de experts worden ingedeeld in drie groepen: (1) experts die veel waarde hechten aan hun privacy en zich (grote) zorgen maken over mogelijk misbruik van hun persoonlijke data, (2) experts die waarde hechten aan hun privacy maar zich niet veel zorgen maken over datamisbruik, en (3) experts die niet veel belang hechten aan hun privacy en weinig doen om hun persoonlijke data te beschermen. Deze drie groepen komen overeen met de categorieën die beschreven zijn in Westin's (1967) Privacy Index Segmentation: privacyfundamentalisten, privacypragmatici en onbezorgde gebruikers. Hoewel de privacyfundamentalisten onder de experts aangeven dat ze bescherming van hun persoonlijke gegevens belangrijk vinden, maken ze onbepert gebruik van online services, zelfs als deze meer permissies vragen dan nodig. De experts rechtvaardigen hun gedrag door te verwijzen naar tijdsgebrek, groepsdruk, gemakzucht of een sterke behoefte om de app te gebruiken. Wel proberen de experts in deze groep de permissies van een app te beoordelen en risico's in te schatten. De privacypragmatici geven aan dat ze zich niet altijd goed voelen als ze online services gebruiken en zijn zich bewust van risico's, maar dit weerhoudt hen er niet van gebruik te maken van online diensten. De onbezorgde experts geven aan dat ze veel afweten van privacyrisico's, maar geen bezwaar hebben om hun persoonlijke informatie online prijs te geven. Het onderzoek bevestigt dat technische kennis over online privacy niet automatisch leidt tot een bewustere omgang met persoonlijke gegevens online. Experts lijken net zo kwetsbaar te zijn op het gebied van online privacy als gewone gebruikers.

Hoofdstuk 5 gaat over verschillen in privacypercepties tussen gebruikers. Met behulp van de Q-sort methode zijn deelnemers ($N = 100$) op basis van hun opvattingen over privacy in groepen ingedeeld. Het onderzoek is opgezet aan de hand van de theorie van contextuele integriteit (Nissenbaum, 2004, 2011), die de hoofdcontext (in dit geval het verschil tussen een gezondheidsapp en een nieuwsapp) en drie contextuele factoren onderscheidt: welke specifieke persoonlijke informatie verzameld wordt (*wat*), hoe deze informatie verwerkt wordt (*hoe*) en welke actoren daarbij betrokken zijn (*wie*). De resultaten laten zien dat de voorkeuren ten aanzien van privacy niet afhankelijk zijn van het type app of de actoren die de data ontvangen en verwerken. De persoonlijke informatie die gevraagd wordt en de manier waarop deze verwerkt wordt blijken daarentegen wel een belangrijke te spelen rol. De resultaten bevestigen de theorie van contextuele integriteit dus maar gedeeltelijk. Voor de twee contextuele factoren

die wel een rol blijken te spelen zijn elk drie gebruikersgroepen geïdentificeerd die verschillen in praktijken die ze (op verschillende gronden) acceptabel vinden. Over praktijken die als serieuze overschrijding van privacygrenzen worden gezien, zijn gebruikers het echter grotendeels eens.

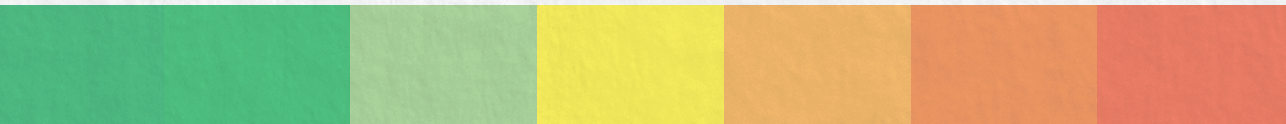
De tweede doelstelling van het proefschrift richt zich op de vraag hoe privacy-aspecten van online services op een visuele en begrijpelijke manier naar de gebruikers kunnen worden gecommuniceerd. Om deze vraag te kunnen beantwoorden is er een tweede literatuuronderzoek uitgevoerd. **Hoofdstuk 6** beschrijft een analyse van bestaande privacyvisualisaties ($N = 13$) en richtlijnen voor 'privacy-by-design' (PbD; $N = 14$). Op grond van het literatuuronderzoek is er een overkoepelende lijst van 15 privacy-attributen opgesteld: *aansprakelijkheid* (accountability), *anonymisering* (anonymisation), *verzameling* (collection), *controle* (control), *juistheid* (correctness), *openbaarmaking* (disclosure), *functionaliteit* (functionality), *pseudonymisering* (pseudonymization), *doeleinde* (purpose), *bewaring* (retention), *recht om te worden vergeten* (right-to-be-forgotten), *verkoop* (sale), *veiligheid* (security), *uitwisseling* (sharing), en *transparantie* (transparency). Met behulp van privacy-experts is deze lijst gevalideerd, verfijnd en zijn er definities voor elk attribuut geformuleerd. Om het belang van de gevonden privacy-attributen te onderzoeken is aan 385 gebruikers en 100 privacy-experts gevraagd de 15 attributen te ordenen op relevantie. Hoewel het doorverkopen van data door gebruikers als belangrijk aspect van online privacy wordt beoordeeld, is dit attribuut weinig of niet in PbD richtlijnen en visualisaties te vinden. Hetzelfde geldt voor de attributen aansprakelijkheid, anonymisering en het recht om te worden vergeten. Geconcludeerd wordt dat bestaande visualisaties niet goed op de behoeften van gebruikers aansluiten. Ook is er nog steeds geen algemeen geaccepteerde privacyvisualisatie ontwikkeld, hoewel dit wordt aanbevolen in de AVG.

Hoofdstuk 7 beschrijft hoe een privacyvisualisatie (de *Privacy Rating*), afgestemd op de behoeften van gebruikers, eruit kan zien. De lijst van privacy-attributen zoals beschreven in hoofdstuk 6 is als basis genomen voor de ontwikkeling van de privacyvisualisatie. Van de 15 aspecten zijn er uiteindelijk 12 meegenomen in de visualisatie. Deze zijn in vier clusters ingedeeld: *Verzameling* (Collection), *Uitwisseling* (Sharing), *Controle* (Control), *Veiligheid* (Security). Vervolgens is elk aspect meetbaar gemaakt op een 3-punt-schaal (goed-neutraal-slecht). Deze schaal correspondeert met strafpunten die aan een online service kunnen worden toegekend (0 strafpunten voor een goede score, 1 punt voor een neutrale score en 2 punten voor een slechte score). Het totaal aantal strafpunten (beste score = 0, slechtste score = 24) werd ingedeeld in zeven categorieën, van A tot G. De privacycategorieën corresponderen met een kleurschema dat loopt

van groen (A) tot rood (G). De visualisatie heeft drie doelen: (1) de visualisatie vestigt de aandacht van gebruikers op het belang van privacy, (2) de visualisatie helpt gebruikers die weinig aandacht hebben voor privacy om zich snel een totaaloordeel te vormen over de privacy-aspecten van een online service, en (3) de visualisatie helpt gebruikers met meer interesse in privacy om gedetailleerde informatie over de privacy-aspecten van een online service te vinden. Uiteindelijk maakt dit ontwerp duidelijk dat het mogelijk is om een overzicht van de meest belangrijke privacy-informatie te maken. De visualisatie is vervolgens getest in een gebruikersonderzoek ($N = 30$). De deelnemers begrepen de *Privacy Rating* goed en vonden het een waardevol initiatief. Daarnaast had de visualisatie een significant effect op het vertrouwen dat de deelnemers in de online service hadden. Alle deelnemers uitten de wens dat de *Privacy Rating* een officiële en vereiste standaard voor online services wordt.

De bovengenoemde studies laten zien dat online privacy veel verschillende nuances en facetten heeft. Het is een concept dat continu in beweging is, naarmate context, persoonlijke behoeftes, voorkeuren en technologieën veranderen. Verder kan geconcludeerd worden dat het prijsgeven van informatie niet op rationele kosten-baten-afwegingen gebaseerd is. Het nemen van beslissingen wordt vaak beïnvloed door cognitieve vertekeningen en heuristische. Ook vindt er in veel gevallen nauwelijks een inschatting van de risico's plaats. Dit geldt niet alleen voor gewone gebruikers maar ook voor gebruikers met bovengemiddelde technische kennis. Het lijkt erop dat gebruikers wel op een algemeen niveau zorgen over privacy kunnen uiten, maar dat deze vaak worden gerelativeerd op het moment dat een online dienst wordt gebruikt. Daardoor komt uiteindelijk een discrepantie tussen attitude en gedrag, zoals beschreven in de privacy paradox, tot stand. Als de aandacht van gebruikers expliciet op het thema online privacy wordt gericht en het bespreekbaar gemaakt wordt, blijken de meeste gebruikers wel controle te willen hebben over hun gegevens. Ze maken dan ook duidelijk dat ze hiervoor ondersteuning nodig hebben. Communiceren over specifieke privacyrisico's kan gebruikers in staat stellen om weloverwogen beslissingen te nemen over hun persoonlijke data. Dit proefschrift demonstreert dat het mogelijk is om met een privacylabel dat de complexe privacy-informatie samenvat, vereenvoudigt en vergelijkbaar maakt, gebruikers te motiveren om meer aandacht te schenken aan hun online privacy. Hierdoor wordt aan het overkoepelende doel van het proefschrift voldaan, namelijk gebruikers in staat te stellen zodat ze weloverwogen beslissingen kunnen nemen met wie ze hun data willen delen en voor welk doel.





Biography

About the author

Susanne Barth was born on the 20th of January 1982 in Herdecke, Germany. Her family left the country's Ruhr Region while she was still young, moving to a small village called Epe in Münsterland, close to the Dutch border.

After finishing her A Levels, she completed her vocational education to become a qualified (with honors) haute-couture tailor in Düsseldorf. Several creative years later, she followed her passion for journalism and moved to Milan, Italy. There, she worked as correspondent and public relations manager at a correspondence office for several German lifestyle and fashion magazines. After four years in Italy, Susanne Barth moved back to Germany, assuming responsibility in Munich for a corporate luxury & lifestyle magazine and a travel guide.

Inspired by years of practical work in the communications field, Susanne Barth decided to pursue an academic path she'd been wanting to follow since her fledgling years in journalism. After determining the right international environment, she moved to Enschede in the Netherlands before starting her bachelor's degree in communication at the University of Twente in 2009. It was during that time that Susanne discovered her keen interest for human behavior, taking on a pre-master program in psychology to compliment her existing studies. In 2012, she not only completed her master's degree in communication with a specialization in marketing communication, but a master's degree in psychology with specializations in human factors and media psychology (cum laude). After her studies, Susanne stayed in the Netherlands where she was a University of Twente, department of Communication Science BMS faculty member. During that time, she lectured in the Communication Science bachelor program. Her other duties included facilitating the re-structuring of the bachelor program in accordance with the newly introduced TOM model and supervising students at both the bachelor and master levels. Later, Susanne was responsible for a module on communication channels and media use - and a module focusing on the Privacy Paradox.

During her teaching activities, Susanne Barth began drawing up the framework for a research project focusing on online privacy. In 2015, she was given the green light to pursue the study, becoming a PhD candidate at the Services and CyberSecurity Group, Faculty of EEMCS and the Communication Science Research Group, Faculty of BMS at the University of Twente. Under the supervision of Prof. Dr. Menno D. T. de Jong and Prof. Dr. Marianne Junger - and the project coordination and advisory of Prof. Dr. Pieter H. Hartel, Susanne started her research within the interdisciplinary SERIOUS research project, funded by NWO and conducted in collaboration with TNO, WODC and Centric.

SERIOUS primarily deals with the privacy and security requirements for mobile applications. During her years as a full-time researcher, Susanne continued teaching, also conducting courses on cybercrime science in the computer science master program. In 2015, Susanne got married and gave birth to a son in 2016.

Currently, Susanne Barth works as a post-doctoral researcher at the Services and CyberSecurity Group, faculty of EEMCS at the University of Twente. Her research focuses on online privacy in general, user perception of online privacy and stakeholder engagement in the development of practical privacy protection measures. Her goal is to make the data handling practices of online services more transparent in order to move toward a balanced relationship between users and online services.

Scientific publications

- Barth, S. & De Jong, M. D. T. (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics* 34(7), 1038-1058. <https://doi.org/10.1016/j.tele.2017.04.013> [ISI; IF 3.714]
- Barth, S., De Jong, M. D. T., & Junger, M. (2020). *Lost in privacy? Online privacy from a cybersecurity expert perspective*. Manuscript submitted for publication.
- Barth, S., De Jong, M. D. T., Ionita, D., & Junger, M. (2020). *Understanding the Android permission system. A cybersecurity expert perspective*. Manuscript in preparation for publication.
- Barth, S., De Jong, M. D. T., Junger, M., Hartel, P. H., & Roppelt, J. C. (2019). Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. *Telematics and Informatics*, 41, 55-69. <https://doi.org/10.1016/j.tele.2019.03.003> [ISI; IF 3.714]
- Barth, S., Hartel, P. H., Junger, M., & Montoya, L. (2019). Teaching empirical social-science research to cybersecurity students: The case of “thinking like a thief”. *IEEE Security & Privacy*, 17(3), 8-16. <https://doi.org/10.1109/MSEC.2018.2888781> [IF 1.596]
- Barth, S., Ionita, D., & Hartel, P. H. (2020). *Understanding online privacy – A systematic review of privacy visualizations and Privacy by Design guidelines*. Manuscript submitted for publication.
- Barth, S., Ionita, D., De Jong, M. D. T., Hartel, P. H., & Junger, M. (in press). Privacy Rating: A user-centered approach for visualizing data handling practices of online services. *IEEE Transactions on Professional Communication*.
- Barth, S., Ngo, T., De Jong, M. D. T., & Krämer, N. C. (2020). *Using the Q-sort method to understand privacy perceptions*. Manuscript submitted for publication.

- Barth, S., Schraagen, J. M., & Schmettow, M. (2015). Network measures for characterising team adaptation processes. *Ergonomics*, 58(8), 1287-1302. <https://doi.org/10.1080/00140139.2015.1009951> [ISI; IF 2.181]
- Barth, S., Van Hoof, J. J., & Beldad, A. (2014). Reading between the lines: A comparison of 480 German and Dutch obituaries. *Omega - Journal of Death and Dying*, 68(2), 161-181. [ISI; IF 1.127]
- De Jong, M. D. T., Harkink, K. M., & Barth, S. (2018). Making green stuff? Effects of corporate greenwashing on consumers. *Journal of Business and Technical Communication*, 32(1), 77-112. <https://doi.org/10.1177/1050651917729863> [IF 0.900]
- Kegel, R. H. P., Barth, S., Klaassen, R., & Wieringa, R. J. (2017). *Computer literacy systematic literature review method*. (CTIT Technical Report Series; No. TR-CTIT-17-05). Centre for Telematics and Information Technology (CTIT). Enschede, The Netherlands: University of Twente.

Conferences and presentations

- Barth, S., Hartel, P. H., Junger, M., & Montoya, L. (2018). *Teaching empirical social-science research to cybersecurity students: The case of "thinking like a thief"*. Paper presented at the ECCA 2018 Environmental Criminology and Crime Analysis Symposium, Elche, Spain.
- Barth, S., & De Jong, M. D. T. (2017). *The privacy paradox – A systematic literature review*. Paper presented at Etmaal van de Communicatiewetenschap 2017, Tilburg, The Netherlands.
- Barth, S. (2017). *The privacy paradox - theoretical and practical implications*. Workshop presented at The International Internet of Things Day 2017, Rotterdam, The Netherlands.
- Carvajal Gallardo, I., & Barth, S. (2015). *Security requirements for serious apps*. Project presented at the NCSRA symposium NWO, 2015, Den Haag, The Netherlands.
- Jansen, M. G. M., Jansma, S. R., Barth, S., & De Jong, M. D. T. (2015). *Push or match? Een vergelijking van informatiebehoefte van burgers en aangeboden informatie van partijen en media*. Paper presented at Etmaal van de Communicatiewetenschap 2015, Antwerp, Belgium.

Media exposure

- Barth, S., (2017, May 16). Uitsmijter Susanne Barth - Een kijkwijzer voor apps. *Computer Idee*, 2017-11.
- Barth, S. (2019, September, 25). Wat is privacy waard? [Privacyweb – Blog]. <https://www.privacy-web.nl/artikelen/wat-is-privacy-waard>

Research project

Sep. 2015 – to date

SERIOUS. SEcurity RequiReiments for SerIOUS apps. *Funded by NWO, the national research council of the Netherlands (Grant number: 628.001.011) in collaboration with TNO (Dutch research institute), WODC (Dutch Research and Documentation Centre) and Centric (Dutch ICT organization).*

