

Schomakers, Eva-Maria; Lidynia, Chantal; Ziefle, Martina

Article — Published Version

All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity

Electronic Markets

Provided in Cooperation with:

Springer Nature

Suggested Citation: Schomakers, Eva-Maria; Lidynia, Chantal; Ziefle, Martina (2020) : All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity, Electronic Markets, ISSN 1422-8890, Springer, Berlin, Heidelberg, Vol. 30, Iss. 3, pp. 649-665, <https://doi.org/10.1007/s12525-020-00404-9>

This Version is available at:

<https://hdl.handle.net/10419/288550>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity

Eva-Maria Schomakers¹ · Chantal Lidynia¹ · Martina Ziefle¹

Received: 15 August 2019 / Accepted: 24 January 2020 / Published online: 5 February 2020
© The Author(s) 2020

Abstract

Privacy-preserving data markets are one approach to restore users' online privacy and informational self-determination and to build reliable data markets for companies and research. We empirically analyze internet users' preferences for privacy in data sharing, combining qualitative and quantitative empirical methods. Study I aimed at uncovering users' mental models of privacy and preferences for data sharing. Study II quantified and confirmed motives, barriers, and conditions for privacy in data markets. Finally, in a conjoint study, trade-offs between decisive attributes that shape the decision to share data are analyzed. Additionally, differences between user groups with high and with low privacy concerns are observed. The results show that the anonymization level has the greatest impact on the willingness to share data, followed by the type of data. Users with higher privacy concerns are less willing to share data in data markets and want more privacy protection. The results contribute to an understanding of how privacy-preserving data markets could be designed to suit users' preferences.

Keywords Data markets · Online privacy · Empirical research · User perspective · Conjoint

Introduction

More and more parts of our lives are *online* and digitally stored as the use of online services has become an integral part of everyday life – be it to search for information, to keep up to date, to stay in touch with friends and family, to work, monitor our health, pass the time, and much more. Thereby, every user creates a large amount of data. The advent of the *Internet of Things* and *Big Data* offers manifold benefits for consumers, businesses, and society. Health care, mobility, production, and education, to name but a few, can considerably profit from the significant data-driven knowledge gain. Personal data become the “new oil” of the digital economy

as it drives innovations, creates knowledge, and allows effectiveness and efficiency in many societally relevant fields (Spiekermann et al. 2015).

At the same time, according to some authors, we have reached the end of privacy (Enserink and Chin 2018). The huge availability of data attracts abusive usage and malpractice. Consumers are worried that they have lost control over their information and report to be highly concerned about their informational privacy (European Commission 2015). In addition, the existing data markets are not transparent: Users are not fully aware of what happens to their data and cannot control data use (Spiekermann and Novotny 2015). A more or less legitimate trade of data in a “shadow market” has evolved (Conger et al. 2013).

This conflict between users' rights and desires for privacy, on the one hand, and market and businesses' demands for data, on the other hand, is still unresolved. One vision that provides a win-win situation for both consumers and businesses are privacy-preserving data markets (Gkatzelis et al. 2015; Matzutt et al. 2017; Spiekermann et al. 2015): A market that provides privacy protection and informational self-determination for consumers and data for business and research, where data are handled transparently and in accordance with customers' consent. And the procedure is open for all of business and research in a viable market that breaks

Responsible Editor: Val Hooper

✉ Eva-Maria Schomakers
schomakers@comm.rwth-aachen.de

Chantal Lidynia
lidynia@comm.rwth-aachen.de

Martina Ziefle
ziefle@comm.rwth-aachen.de

¹ RWTH Aachen University, Campus-Boulevard 57,
52074 Aachen, Germany

up existing monopolies. When willingly shared, when the information by customers is accessible with a low market entry barrier, higher legal certainty, a high transparency for the parties involved as well as good data quality, then the foundation for societally relevant innovations is laid.

In privacy-preserving data markets, users shall actively and voluntarily share their data. The willingness to share data is dependent on many factors, e.g., privacy concerns, benefits, type of information, and culture (Hallam and Zanella 2017; Markos et al. 2017; Trepte et al. 2017). This willingness to share has been extensively studied (Smith et al. 2011) and some studies were conducted within the context of data markets and included privacy protection in data sharing (Roeber et al. 2015; Ziefle et al. 2016). But we do not yet have the necessary insights into what informational self-determination and privacy means to users in such data markets. How can users' desires for privacy be satisfied while they share data in the decontextualized online world? And how do users form decisions regarding data sharing in such a new environment? Because of this knowledge gap, we empirically analyze users' notions of and desires for privacy protection in a privacy-preserving data market.

We use a threefold mixed-method approach to identify and understand users' notions of privacy in data sharing in detail as well as to quantify users wishes and observe trade-offs in data sharing decisions. Therefore, we start with a qualitative assessment of users' understanding and wishes for privacy, in which underlying mental models, attitudes, and visions about privacy, data sharing, and the conditions for it are identified. Based on the findings of this first study, users' motives, barriers, and conditions for privacy in data sharing are then measured and quantified in a second study (survey). The third step incorporates a choice-based conjoint study, a method that is very valuable to determine the importance of and trade-offs between different attributes in realistic decision scenarios which are a combination of conflicting factors. The participants are asked to decide in different simulated scenarios if they would be willing to share data as to understand which factors are decisive in users' decision making for data sharing. This triangulation of empirical methods enables us to combine the benefits of each method and thus to answer our research question holistically.

The following section gives an overview over the theoretical approach and related work to users' online privacy behaviors and attitudes before each study's methodology and results are presented.

Related work: Online privacy and data sharing from the user perspective

As Nissenbaum describes, researchers can agree on one fact: privacy is messy, complex, and hard to define (Nissenbaum

2010). Nonetheless, a description of online privacy from a user perspective is approximated in the next section and related work regarding users' attitudes and public perceptions of online privacy and data sharing are briefly outlined.

Online privacy and data sharing

As one essential work, Altman defines privacy as “the selective control of access to the self” (Altman 1976, p. 8) that requires an active and dynamic regulatory process. As such, the desired state is not to keep everything secret and to withdraw completely from people but to find the right balance between withdrawal and disclosure in the given context. For that, the control over “when, how, and to what extent information about [oneself] is communicated to others” is decisive (Westin 1967, p. 7). The concept of *informational self-determination*, first established by the German Bundesverfassungsgericht (Federal Constitutional Court) as a basic constitutional right in 1983, describes exactly this active and cognizant empowerment of users (Hornung and Schnabel 2009).

But in the online world, users feel to have lost control (European Commission 2015). Disclosed information is persistently available over space and time, is searchable, shareable, and replicable (Palen and Dourish 2003; Taddicken 2014). From early childhood on, we learn that we can manage our privacy by means of physical space and visible environments: we close doors and curtains, lock doors, and hold conversations away from unwanted listeners. In the digital world, we cannot rely on these natural and intuitive mechanisms anymore. The desired state of privacy, and with this the desired sharing of information, is defined by context and audience (Nissenbaum 2010). Online, contexts collapse and audiences become heterogeneous (Marwick and Boyd 2011; Taddicken 2014). Even malicious acts targeting our privacy become more harmful and far-reaching. Aggravatingly, users might not know and might not be able to realistically assess the harmful situation, as the danger of malpractice is not necessarily visible. Many internet users perceive that they have lost control over their information and they are highly concerned about their online privacy (European Commission 2015; Li 2011). 58% of Europeans see no alternative to the provision of personal information to obtain products or services (ibid.). Active privacy protection which could soothe concerns is perceived as too complex, not feasible, and users are resigned (Hoffmann et al. 2016). On the other hand, protection is neglected because of a feeling of “nothing to hide” (Prettyman et al. 2015) or “no-one is interested in my data” (Schomakers et al. 2018).

This feeling of being overpowered and the loss of control experienced by the users could be one explanation for the so-called *privacy paradox*. A number of studies have found a mismatch between users' privacy attitudes and their respective

online behavior: Users are concerned about their online privacy, but they still share private information without hesitation and do not protect their privacy accordingly (Gerber et al. 2018; Kokolakis 2015). But research has also yielded opposing results: Users are willing to pay extra for privacy and concerns indeed trigger protective responses (Baruh et al. 2017; Blank et al. 2014; Egelman and Felt 2012; Lutz and Strathoff 2013). The *privacy calculus* theory postulates that users weigh the anticipated risks for their privacy against the perceived benefits of data disclosure (Dinev and Hart 2006). But users cannot foresee all of the consequences and risks of said disclosure. On the one hand, humans' cognitive capacity and complexity skills are limited and decisions are therefore based on heuristics and biases formed by experience (Acquisti et al. 2015). On the other hand, even if we could process the complexity of information, we do not have full knowledge about what may happen to our data in the present and future (Baruh and Popescu 2017; Spiekermann and Novotny 2015).

In the vision of a privacy-preserving data market, users are provided with full informational self-determination. In contrast to data repositories maintained by digital companies, who collect information about their customers, and existing third-party data markets, in privacy-preserving data markets users are not only informed about proper data handling, data are actively distributed by the users to the privacy-preserving data market, thereby enabling full control over the information flow and privacy (Matzutt et al. 2017). This also adheres to the demands of the new European General Data Protection Regulation (GDPR) that has come into force in 2018. This paper aims to understand which privacy protective measures and which control possibilities users desire to maintain their privacy when sharing data.

Privacy protection in data sharing

To interact with complex technical systems, people employ mental models. Mental models are cognitive representations of how technical systems or interfaces might work, including persons' beliefs as well as cognitive and affective expectations about their functions (Gentner and Stevens 2014). Mental models are not necessarily correct representations of the real functioning of the real systems, but still helpful in guiding users and helping them understand and interact with complex and abstract issues and systems (Asgharpour et al. 2007; Morgan et al. 2002).

Mental models concerning perceived risks in the online environment and online privacy are often transferred from the physical world. Five predominant mental models of users regarding online threats have been identified in previous research (Camp 2009): physical security (e.g., break-in in your home), medical security (e.g., computer virus), criminal behavior (e.g., identity theft, vandalism), warfare (threats have fast response times and huge potential losses), and economic

failure (financial losses). Especially physical security – meaning security from physical entrance to your private space(s) and belongings – are used to describe the risks of the online environment (Asgharpour et al. 2007). Security is perceived as a barrier, therefore items locking someone out or preventing access to information are associated with privacy and security, e.g., locks, keys, and shields (Dourish et al. 2003; Motti and Caine 2016).

These real world concepts that users take to make sense of the virtual world – the quintessential idea of mental models (Asgharpour et al. 2007; Coopamootoo and Groß 2014) – restrict access and data sharing completely. So how can these concepts be translated and applied to the context of data sharing in a privacy-preserving data market? One possibility to protect users' privacy in data markets is anonymization of the shared data. The level of anonymity can be described and acquired in different ways. One concept is the *k*-anonymity. If a person has *k*-anonymity within a data set, this person is not distinguishable from at least *k*-1 other individuals in this data set. Therefore, with $k > 1$, the person is not 100% identifiable.

In real-world decisions, often conflicting aspects have to be considered and weighed up – like benefits and barriers are weighed against each other according to the Privacy Calculus. In the trade-offs between positive and negative aspects, the relevance of each factor for the individual can be revealed. To do so empirically, Conjoint studies can be conducted. Prior Conjoint studies included different as well as similar factors and show partly contradicting results. In Online Social Networks (OSN), it could be shown that free of charge access compared to small payments had the most impact on users' data sharing decisions. Privacy control mechanisms for the visibility of posts to other users, on the other hand, were less important (Krasnova et al. 2009). For data sharing with organizations, Roeber et al. (2015) found that the monetary compensation was the second most important factor and how the data is used by organization (including anonymized use as options) again less important. Here, the most important factor for users' decision to share data was the data type. For sharing medical data, two anonymization techniques, *k*-anonymity and differential privacy, have been compared in their impact on sharing decisions by Calero Valdez and Ziefle (2019). Regardless of the anonymization method, anonymization was the most important factor for the participants in this study and the type of benefit for data sharing was less important. These conflicting results can be contributed to different reasons: For one, different sample and cultural and contextual settings for data sharing were used in these studies. For another, the selection and exact operationalizing of the factors, as well as the number of levels and level selection influence the results of Conjoint studies extremely. Therefore, the difference in reward from 0 to 50€ in the study by Roeber et al. (2015) to a not further specified

type of benefit being global, financial, or personal (Calero Valdez and Ziefle 2019) makes a huge difference and makes these studies theoretically incomparable.

Still, the qualitative and quantitative research approaches to understand privacy in data sharing have not yet been brought together. There is a need for a holistic approach to reveal users' notions about anonymization, their understanding of this abstract concept, and wishes for anonymization levels as well as to show how users attribute relevance to privacy protection in trade-off with other important factors. Prior Conjoint studies have used specific context (e.g., OSNs, medical data sharing) and different operationalization of privacy in data sharing (e.g., privacy control for visibility of posts in OSNs, binary levels of anonymized data compared to data linked to the name). To understand privacy in data sharing in privacy-preserving data markets, research within this specific context considering the users' mental models of privacy protection is needed.

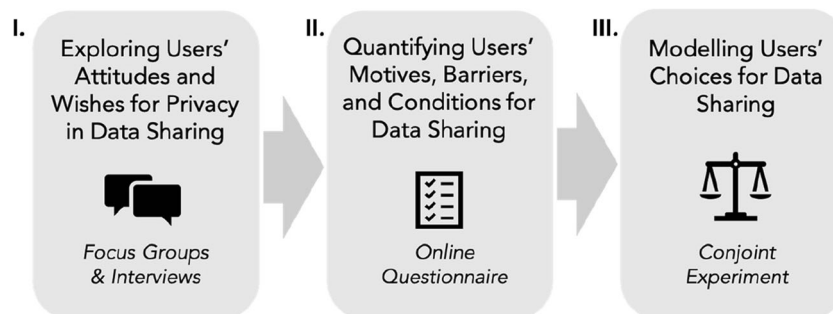
Material and methods: Empirical approach

In order to understand users' notions of privacy in data sharing, we undertook a three-step empirical approach. In a first exploratory approach, qualitative interviews and focus groups were used to uncover mental models and attitudes regarding privacy in data sharing and conditions for a user-centered privacy-preserving data market. Based on these findings, an online questionnaire was developed to quantify users' motives, barriers, and conditions as well as willingness to provide data using a larger sample of German internet users. In a third-step, experiments were conducted to model users' choices of data sharing in which the trade-offs between privacy protection, type of data, data receiver, and benefit of data sharing are identified. Figure 1 illustrates our approach.

Study I: Exploring users' understanding of privacy in data sharing

The aim of Study I was to examine users' notions regarding privacy in data sharing and to identify motives, barriers, and

Fig. 1 Three-tier research approach



conditions for the use of user-centered privacy-preserving data markets. A qualitative approach was chosen, as it is not yet known what users' preferences for such data markets are and especially mental models regarding privacy usually do not include the case of sharing data.

The research process of study I

In total, 7 interviews and 6 focus groups (guided group discussions) were conducted. All followed a similar interview guideline (cf. Figure 2) but put different foci on the topics. After an introduction and warming up with discussions about privacy in the online environment, the scenario of a user-centered, privacy-preserving data market was introduced. The participants were asked, what privacy protection means to them when sharing data. Three focus groups put focus onto this topic and included further exercises to elicit mental models about privacy in more detail (the preliminary results of these are published in Schomakers et al. 2018). In all focus groups, the participants were not only asked to *talk* about their notions of privacy in such a data market but also to *draw* visualizations and controls for privacy in data sharing. The second topic dealt with motives to use a privacy-preserving data market and barriers which hinder usage. Additionally, conditions (technical, social, economic and more) were discussed.

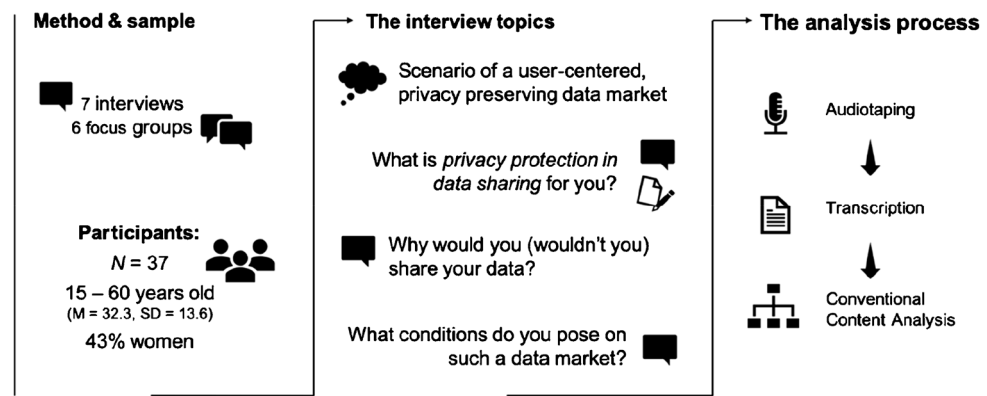
The interviews and focus groups were audiotaped and transcribed verbatim, before categories were derived from the data following a conventional content analysis (cf. Hsieh and Shannan 2005).

37 people between the age of 15 and 60 years ($M = 32.3$, $SD = 13.6$) participated. 43% were women. The participants were recruited with the goal to include people of differing level of privacy knowledge and digital skills, age, and gender.

Results of study I

The reporting of the results is structured as follows: first, the ideas and concepts regarding privacy protection are presented before attitudes about and conditions for data sharing are described.

Fig. 2 Method and sample of study I



Concepts of privacy protection

As taken from the users' reports, privacy protection in general is not a well-defined concept: It appears as multi-layered, contextualized, and participants differ heavily in their understanding of the topic. "Exhausting," "complex," even "impossible" but still very "important" were the attributes used for privacy protection in general. Privacy protection is perceived as multi-layered and contextualized as there are various parties of responsibility for privacy protection, different perceived threats, and, thus, different options for protection. Threats mentioned by the participants are especially criminal behaviors and illicit use of data, but also the collection of data itself and resulting targeted advertising. The demand for more controllability and self-determination is plain. Moreover, some participants saw a threat for society and democracy, e.g., because of manipulation possibilities and filter bubbles which may influence public opinion and elections.

Analyzing the perceived threats when participating in the online world is a key approach to understand mental models of online privacy protection, as some participants defined privacy protection explicitly as the absence of negative consequences. One conceptual threat was implicitly addressed very often throughout the discussion and interviews: *identification*. Participants oppose the notion that others can "form a picture of themselves" (a German idiom meaning to form an opinion about themselves), that "apps can identify yourself," that "data are traced back to yourself," and that "data are combined to individual profiles." Central is the identity and the "me" that participants want to protect. Their wordings and statements suggest that they do see identification as the key problem and not being identified as the one mechanism that can protect them online. Focusing on the idea of sharing data while preserving privacy, the control over what is shared to whom for what benefit and which purpose is one central element. The other is guaranteeing anonymity (see the following quote).

"The most important thing for me would be anonymity"
(male participant, 29 years old)

Notions of privacy protection are mostly 'binary': privacy is either protected or not. This matches most learned, real world concepts of privacy protection, e.g., the use of locks, shutting doors, or building a fence. Still, the idea of gradual anonymity can be learned. Here, the concept of k-anonymity can be easily understood and applied to the idea of privacy-preserving data markets.

Motives and barriers for sharing data in a privacy-preserving data market

As motives for using such a data market mostly benefits for oneself or others were discussed. Benefits for oneself do not only include gratification (e.g., monetary) but also to get an overview over one's data and online accounts. Benefits for others are especially seen regarding the use of the shared data for science and medicine. Here, users are more willing to share data for such a benefit. The feeling of being in control when the data market puts informational self-determination into practice is another decisive driver.

"If the data would only be used for medical purposes, for improvements in science or so, then I would be more willing to provide data." (female participant, 51 years old)

Perceived barriers include essentially three topics: privacy concerns and missing trust into the security of such a data market, moral concerns, and not seeing the personal benefit of the system. Privacy concerns were quite foreground in the discussions and range from fears of the provider of the data market being dishonest, the security of the data market being insufficient, to the data receivers abusing the data material, e.g., with using it for different purposes than consented to. Some participants stated that they would never use such a data market because they morally disapprove with 'selling' data. Others saw "no benefit at all" in the concept (male participant, 26 years old).

“I don’t trust it. Even if the provider would be honest, how will he guarantee that the data buyers use it only for the stated purposes, or that no hackers can get access.” (female participant, 21 years old)

Conditions for data sharing

The participants addressed multiple factors that influence their willingness to provide data. Of great importance were the data receiver and the purpose of the collection. Here, the benefits for the self or the society, e.g., for medical studies, are one important aspect. Also, when the participants can understand why the data are useful to the receiver, they seem to be more willing to provide data. One huge point for discussion was, what types of data are shared and their perceived sensitivity. Of additional importance for the users are the questions how the data are collected and stored, how long it is stored, the prevailing options to let personal data be deleted, and transparency. For privacy protection, the participants demanded – besides informational self-determination and transparency – anonymity and protection within the data market application (e.g., password protection, facial recognition).

All in all, the participants’ attitudes towards user-centered privacy-preserving data markets were quite diverse. Some participants showed very high privacy concerns in general and regarding such data markets, while others were more laid-back. Also, some participants saw moral barriers against trading data, where others were focused more on the personal benefit.

Key Findings of Study I

- A key element of online privacy and privacy in data sharing is the protection of the identity, and thus, anonymity.
- For privacy in data sharing, it is important to control.
 - what types of data are shared (and as how sensitive these are perceived),
 - with whom (and how much the person is trusted),

- for what purposes,
- and for which benefit.
- Barriers against the use of privacy-preserving data markets are foremost privacy concerns, as well as moral concerns against the ‘trading’ with personal data.
- Users’ are diverse regarding their attitudes, especially they differ regarding the strength of their privacy concerns.

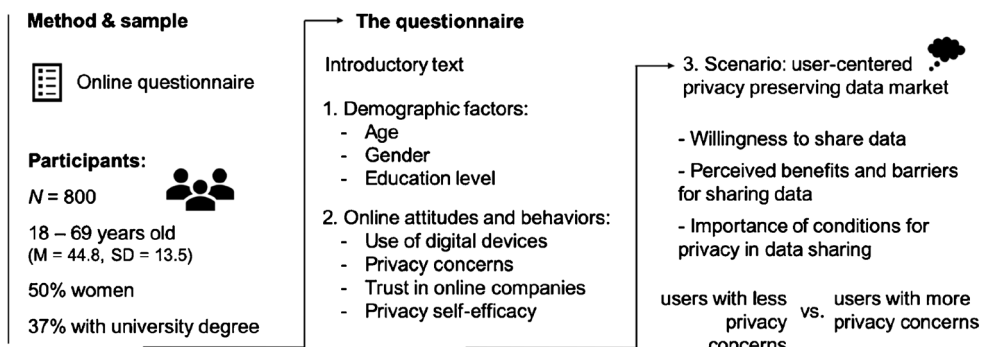
Study II: Quantifying users’ conditions for data sharing

The second study aimed at confirming and quantifying the motives, barriers, and conditions for data sharing identified in Study I. $N = 800$ German participants completed an online questionnaire in which they were introduced to the scenario of a privacy-preserving data market. In Study I, varying preferences towards such data markets could be observed between people with less and more privacy concerns. Therefore, in Study II, we also analyze differences between groups of users with lower and higher privacy concerns.

The research process of study II

The online questionnaire consisted of three main parts (cf., Fig. 3). The third and main part of the questionnaire started with a scenario of a privacy-preserving data market: A newly developed data market enables participants to share their data with interested parties and to get rewards in return. This application values the users’ privacy: Data are shared anonymously, the users can control exactly what information is shared with whom, and they get full transparency. One of four scenarios was randomly assigned to each participant. The scenario differed only regarding the context and type of data: smart home data, customer service data, medical data for clinical trials, and data about online behavior for market research. After the introduction to the scenario, the participants indicated their willingness to share exemplary data types, and their agreement to benefits, barriers, and conditions for the use of

Fig. 3 Description of the online questionnaire of study II



the data market. Beforehand, demographic factors and online attitudes and behaviors were assessed (items for privacy concern based on Xu et al. (2008), trust based on McKnight et al. (2002), privacy self-efficacy based on Schomakers et al. 2019a). All items were measured on symmetric 6-point Likert scales ranging from 1 ('I do not agree') to 6 ('I fully agree'). The reliability of the scales was confirmed using Cronbach's Alpha (criteria $\alpha > .7$ was met by all scales).

For the analysis of group differences, a median split of privacy concern was used. By this, a group of 429 participants with high privacy concerns ($M = 4.79$, $SD = 0.59$, $min = 4.00$, $max = 6.00$) and a group of 371 participants with low privacy concerns ($M = 3.3$, $SD = 0.51$, $min = 1.29$, $max = 3.86$) were compared. The data were analyzed using multivariate procedures. Analyses of variance (Pillai's Trace) were calculated to study differences between the groups. The significance level was set at 5%.

The questionnaire was distributed online via an independent market research company. $N = 800$ Germans participated who are regular internet users. Details of the demographic characteristics of the sample and each group can be taken from Table 1.

Results of study II

The first part of the analysis examines whether the data types differ regarding how willing the participants are to share data in the presented scenario. Figure 4 illustrates the mean willingness to provide data, which differs significantly between the data types ($F(9, 791) = 84.2$, $p < .001$, partial $\eta^2 = .49$). Age is the only data type, that the participants are, on average, willing to provide ($M = 3.84$). Location data ($M = 2.53$) and income level ($M = 2.54$) are most rejected to be provided. Also, activity data, real time data, consumer habits, and even zip-codes, shopping interests, and preferences for technology are seen critically (scores below the midpoint of the scale).

The willingness to provide data differs significantly between user with low privacy concern and with high privacy concern ($F(10,789) = 5.10$, $p < .001$, partial $\eta^2 = .06$): Users

with lower privacy concerns are more willing to provide data than users with higher privacy concerns.

Motives and barriers

Secondly, the motives and barriers regarding the use of a privacy-preserving data market are studied. The agreement to the motives is all in all rather neutral (around the midpoint of the scale of 3.5). Users see mostly benefits in the informational self-determination, privacy protection, and the rewards for data provision (cf. Fig. 5). Trust in data protection and a laid-back attitude towards data collection are rather disagreed to. Decisive barriers to use a user-centered privacy-preserving data market are foremost concerns about privacy and data security (cf. Fig. 6). Additionally, moral arguments and not seeing personal benefits are perceived barriers.

Users with lower privacy concerns see significantly more motives to use the privacy-preserving data market than do users with higher privacy concerns ($F(10,789) = 9.24$, $p < .001$, partial $\eta^2 = .11$). Especially, the trust in the data protection in such a data market is higher by those with lower privacy concerns. Additionally, those with lower privacy concerns do not perceive that many barriers ($F(15,784) = 12.7$, $p < .001$, partial $\eta^2 = .2$). These differences are prevalent for all barriers and all motives.

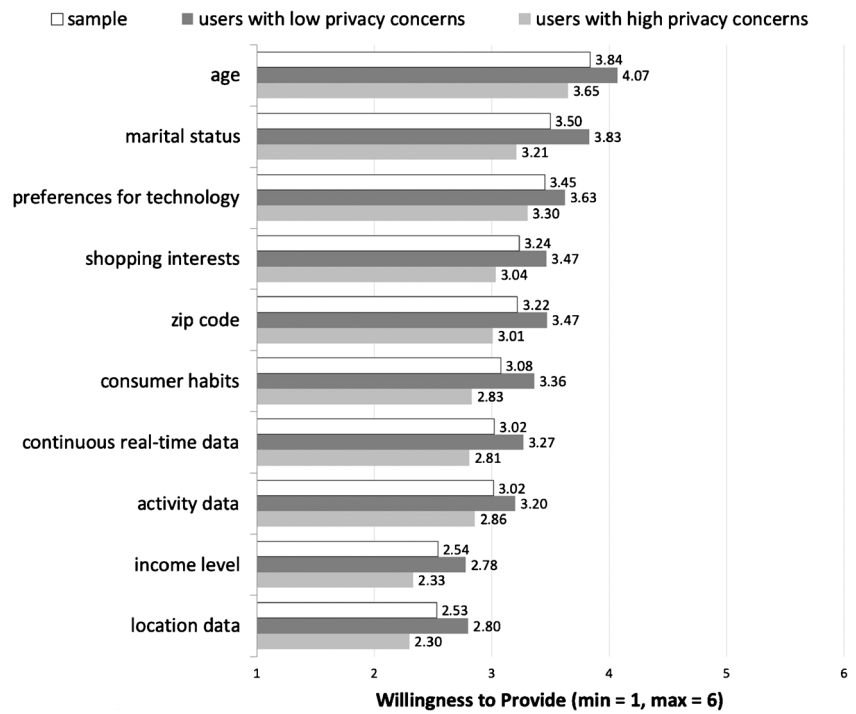
Conditions

The mean evaluations of importance of the conditions are depicted in Fig. 7. All conditions are rated to be between important to very important, with data security ($M = 5.4$, $SD = 0.9$), simple ways of deleting data ($M = 5.34$, $SD = 1.03$), and anonymity ($M = 5.33$, $SD = 1.07$) being the most important. Sufficient rewards ($M = 4.33$, $SD = 1.53$) are least important. Users' with high privacy concerns put more importance on these conditions for privacy than do users with lower privacy concerns ($F(10,789) = 18.13$, $p < .001$, partial $\eta^2 = .19$). These user groups differ in the perceived importance of all conditions with exception of sufficient rewards.

Table 1 Demographic characteristics of the sample and the two concern groups

		Complete sample (n = 800)	Low privacy concern (n = 371)	High privacy concern (n = 429)
age	mean (SD)	44.8 (13.5)	43.22 (13.7)	46.17 (13.2)
gender	women	49.6%	47.4%	51.5%
	men	50.4%	52.6%	48.5%
education	no certificate	1.1%	1.3%	0.7%
	secondary education	15.1%	14.8%	15.2%
	apprenticeship	25.9%	27.5%	24.5%
	qualification for university entrance	20.9%	20.8%	21%
	university degree	37.3%	35.6%	38.7%

Fig. 4 Mean willingness to provide data in a user-centered privacy-preserving data market in comparison between the groups of low and high privacy concern ($n = 800$)



Key Findings of Study II.

- Significant differences in the willingness to provide different data types could be observed between the data types → the type of data to be shared is a decisive factor.
- Informational self-determination and benefits are the most important motives to use a user-centered privacy-preserving data market.
- Privacy concerns are the most relevant barriers against such a data market, followed by moral concerns.

Fig. 5 Mean agreement to the motives of using a user-centered privacy-preserving data market: “I would use the data market because...” ($n = 800$)

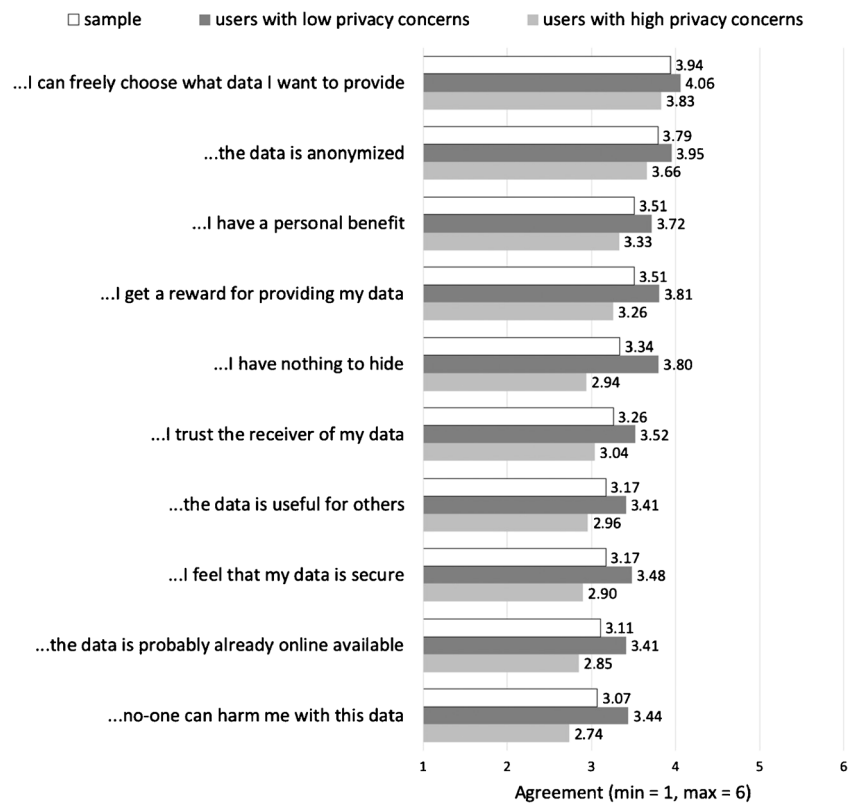
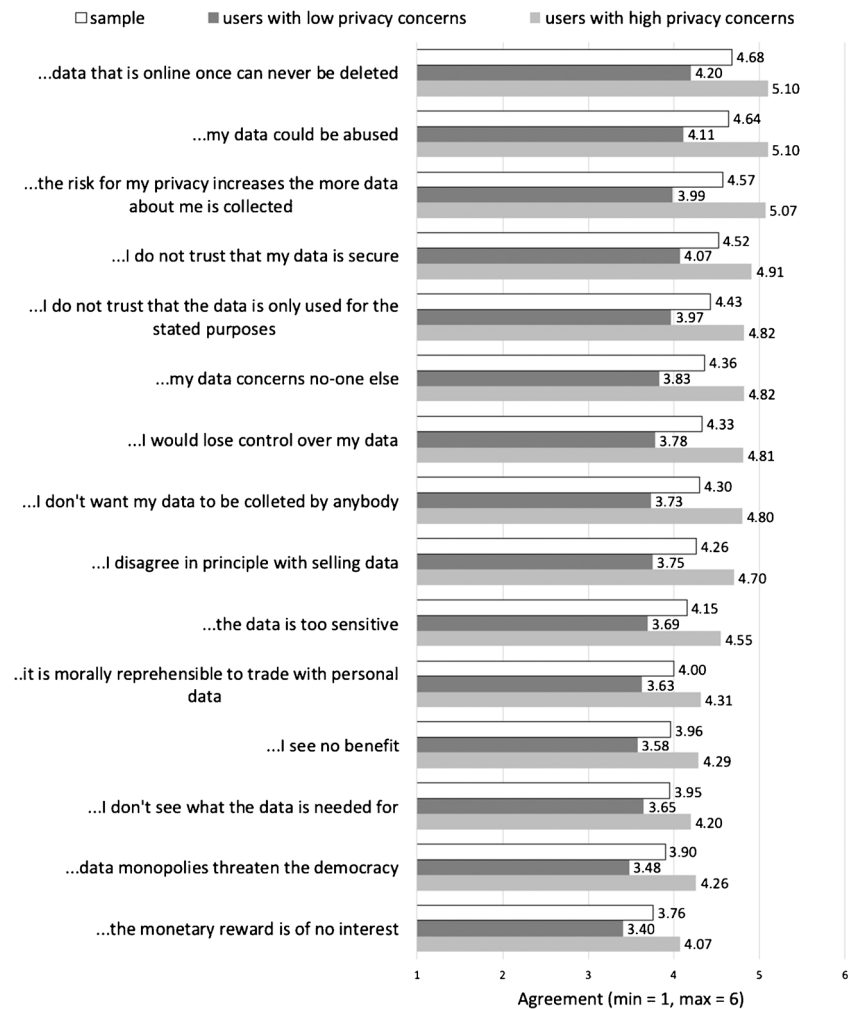


Fig. 6 Mean agreement to the barriers of a user-centred privacy-preserving data market: "I would not use the data market because..." ($n = 800$)



- All conditions for data sharing identified in the qualitative approach in Study I were confirmed to be highly important and differences between conditions were small.
- Users with higher privacy concerns in general are less willing to provide data in a privacy-preserving data market, agree less to motives and more to barriers, and put more emphasize on privacy conditions.

Study III: Modeling users' trade-offs in data sharing

Study II showed that all barriers and conditions for privacy in data sharing were important when asked independently. However, in real world usage scenarios, factors are not independent of each other but might influence and compensate each other, resulting in a weighing of motives, barriers, and conditions. In order to examine which factors are decisive, conjoint experiments were developed for Study III that show the trade-offs between the different factors. As differences

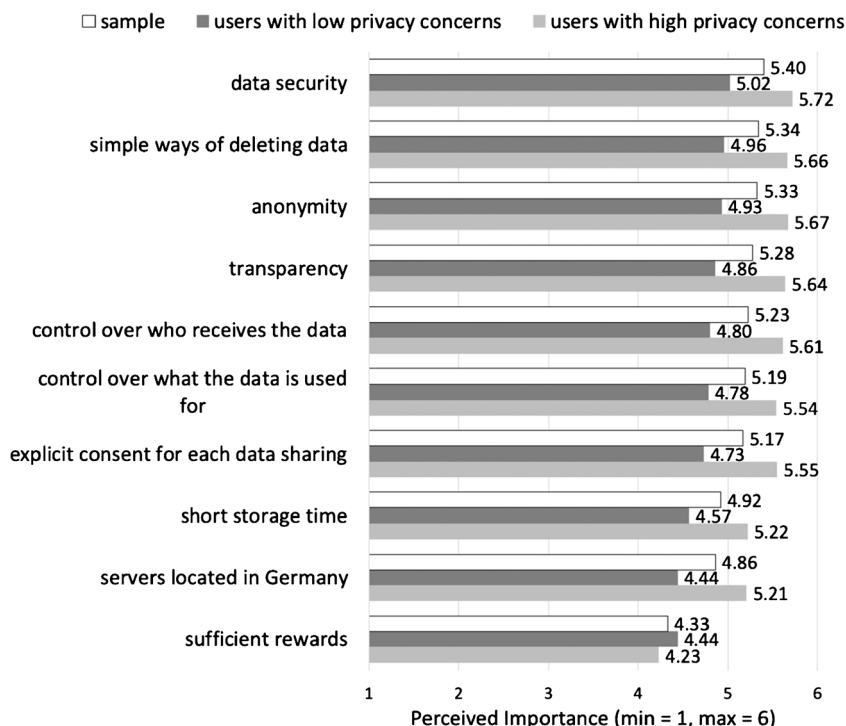
between users with higher and lower privacy concerns could be shown to be relevant, we again look at differences between these groups.

The research process of study III

In reality, decisions are based on different factors in combination, weighing them against each other (Luce and Tukey 1964). In the previous studies, the privacy protection, type of data, data receiver, and the benefit were identified as decisive factors for the decision to share data in a privacy-preserving data market. In study III, these factors are now evaluated jointly in a conjoint experiment to mimic realistic user decisions.

Choice-based conjoint experiments model these complex decision processes: Participants need to choose between n scenarios with m different attributes which vary in levels. They need to take all attributes into account at the same time. As a result of the conjoint analysis, the *relative importance* of attributes shows how much each attribute influences the

Fig. 7 Importance ratings of the conditions for a user-centered privacy-preserving data market ($n = 800$)



decision, and *part-worth utilities* reflect which attribute levels are most and least accepted (Arning 2017).

The attributes and attribute levels

The selection of the attributes for this study was based on the findings in Studies I and II, referring to which factors mostly influence users' decisions to share data. Privacy protection was operationalized in two ways: on the one hand, the level of anonymization was included as k -anonymity; on the other hand, the type of app security was assessed. The type of app security was included as it was shown in Study I to be a typical representation for privacy protection in laypersons' perceptions. Table 2 shows all attributes and attribute levels. Detailed explanations were given to the participants and the anonymization level was additionally instructed with an illustrative example to clarify k -anonymity. Again, here the laypersons' understanding of anonymity was considered: As most participants of Study I showed to have a notion of a 'complete' anonymization, this was included as level in this study. Thereby, the two other presented anonymization levels can be compared to this fictitious and unrealistic threshold that is still a mental model for many users not familiar with technical anonymization capabilities.

Pictograms were used to improve comprehensibility and ease of use. The participants were given 10 random choice tasks, as a complete design with 1024 ($4 \times 4 \times 4 \times 4 \times 4$) possible combinations of the attribute levels would not have been feasible. Tests of design efficiency confirmed the applicability of this reduction. In each choice task, the participants

were presented with $n = 3$ scenarios – that are random combinations of the attribute levels – plus a "none" option. No restrictions in level combination were given and participants were not able to skip tasks.

Questionnaire and sample

In this study, the choices between scenarios were provided within an online questionnaire that started with questions about demographics, usage and experience with apps, and privacy concerns, before the conjoint tasks were introduced (cf., Fig. 8). For the conception and analysis of the conjoint, Sawtooth Software was used.

$N = 126$ complete sets of answers are analyzed. The participants were recruited online, volunteered to take part, and are rather young on average ($M = 27.87$, $SD = 7.9$). Gender was quite balanced with 56% women (cf. Table 3). All participants reported to be frequent internet and app users. For the analysis of group differences regarding privacy concerns, again a median split was used resulting in a group of 54 participants with low privacy concerns ($M = 2.8$, $SD = 0.75$, $min = 1$, $max = 3.75$) and a group of 72 participants with higher privacy concerns ($M = 4.81$, $SD = 64$, $min = 4$, $max = 6$).

Results of Study III

To identify which attribute has the greatest impact on users' decision to share data, *relative importance scores* are calculated. *Part-worth utilities* (zero-centered diffs, calculated using

Table 2 Attributes and attribute levels of the conjoint experiment

benefit				
data type	 location data	 medical history	 social network profile	 financial account data
level of anonymization	 'complete' anonymization	 1 of 5	 1 of 2	 no anonymization
data receiver	 financial institution	 public administration	 non-governmental organization	 online company
app security	 password protection	 double password protection	 fingerprint scanner	 facial recognition

Hierarchical Bayes multinomial logit model) are analyzed to observe how the attribute levels are accepted in comparison to each other. For the interpretation of part-worth utilities, it is important to bear in mind that they cannot be compared between attributes (Gustafsson et al. 2007). Also, positive utilities do not signify acceptance nor do negative utilities signify rejection. The differences between utilities show the influence on the sharing decision in comparison to each other.

Figure 9 pictures the relative importance scores of the five attributes. The *anonymization level* has the greatest impact on the sharing decision for both groups ($M = 34.8\%$). Second most important is the *type of data* ($M = 23.1\%$). *Benefit*, *data receiver*, and *app security* show a similar, smaller influence ($M \approx 14\%$).

The concern groups show small but significant differences in their importance attribution ($F(4,121) = 3.14, p < .05$, partial $\eta^2 = .09$). The relative importance of *data type*, *data receiver*, and *app security* does not differ. But users with a higher privacy concerns put more value to the *level of*

anonymization ($M = 38\%$) than do those with lower privacy concerns ($M = 30.6\%$, $F(1,124) = 12.37, p < .01$, partial $\eta^2 = .09$). Instead, those with lower privacy concerns put more importance to the *monetary benefit* ($M = 15.2\%$) than those with higher privacy concerns ($M = 12.5\%$, $F(1,124) = 4.86, p < .05$, partial $\eta^2 = .02$).

In Fig. 10, the part-worth utilities for all attribute levels are presented. Not surprisingly, higher privacy protection with *'complete' anonymization* is most accepted by the participants whereas *no anonymization* is least accepted. Users rather share their *social network profile* and *location data* than their *medical history* and least their *financial account data*. Also not surprisingly, the higher the monetary reward is, the more are users willing to provide data. As data receivers, *public administration* and *non-governmental organization* are more accepted than *online company* and *financial institution*. The most accepted app security option is *double password protection* followed by *password protection*. *Facial recognition* is the least accepted.

Fig. 8 Description of the online questionnaire and conjoint tasks in study III

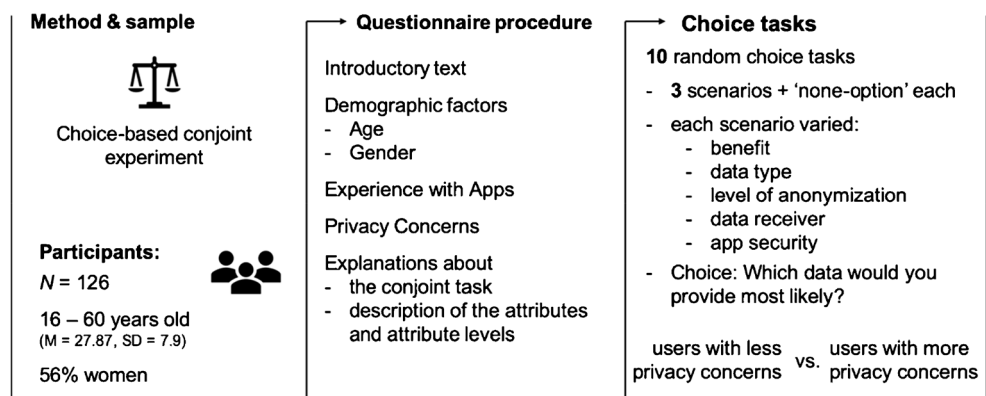


Table 3 Demographic characteristics of the sample and the two groups

		Complete sample (n = 126)	Low privacy concern (n = 54)	High privacy concern (n = 72)
age	mean (SD)	27.87 (7.9)	26.83 (5.9)	28.64 (9.03)
gender	women	56.3%	48.1%	62.5%
	men	43.7%	51.9%	37.5%
education	no certificate	1.6%	0%	2.8%
	secondary education	18.3%	16.7%	19.5%
	apprenticeship	39.7%	37%	41.7%
	qualification for university entrance	27%	24.1%	29.2%
	university degree	13.5%	22.3%	7%

No significant differences in the level evaluation between the groups with low and high privacy concerns can be observed ($F(15,110) = 1.71, p > 0.5$).

Key Findings of Study III

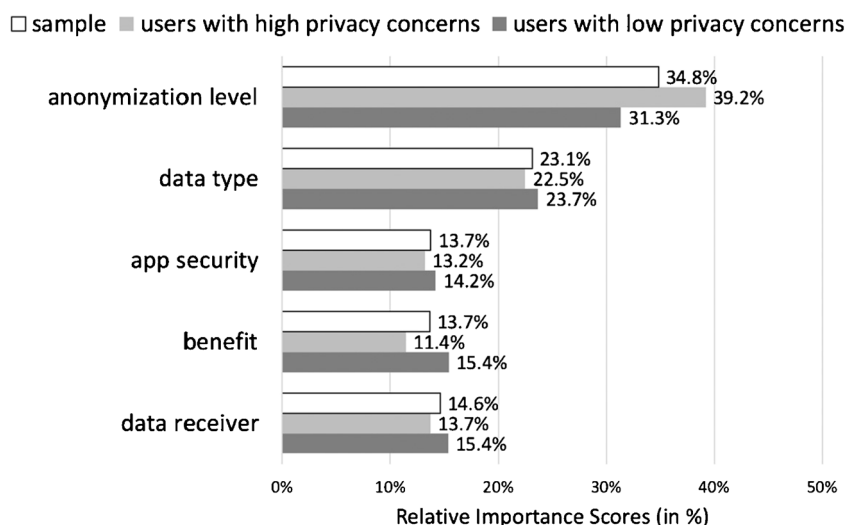
- The *level of anonymization* has the greatest impact on the decision whether to share data, followed by the *type of data*.
- The most accepted scenario is sharing the *social network profile* with ‘complete’ anonymization to a *public administration* for 75€. The application should be protected by a *double password protection*.
- The least accepted scenario is sharing *financial account data without anonymization* to a *financial institution* for 5 € with the application being protected with *facial recognition*.
- Users with high privacy concerns attribute more importance to the anonymization level and less importance to the monetary benefit than users with lower privacy concerns.

Discussion and conclusion

In this paper, we empirically analyzed users’ attitudes and preferences for privacy in user-centered privacy-preserving data markets. These markets are one approach to restore users’ online privacy and informational self-determination as well as to build a viable data market for companies and research. Here, users can decide what data they want to provide to whom for what purpose – and their data are anonymized.

Our research was guided by the question of what users want in order to perceive privacy and to have self-determination in sharing data. A three-fold research approach was used that started with a qualitative study identifying mental models, preferences and relevant conditions. Building on the results, Study II assessed the importance of these motives, barriers, and conditions, as well as the willingness to provide different data types. In a choice-based conjoint experiment, Study III, five attributes (anonymization level, type of data, data receiver, monetary benefit, app security) were weighed against each other in their impact on the users’ decisions to share data. Figure 11 summarizes the key findings. The contribution of this paper is the mixed-method approach to the

Fig. 9 Relative importance scores of the attributes (n = 126)



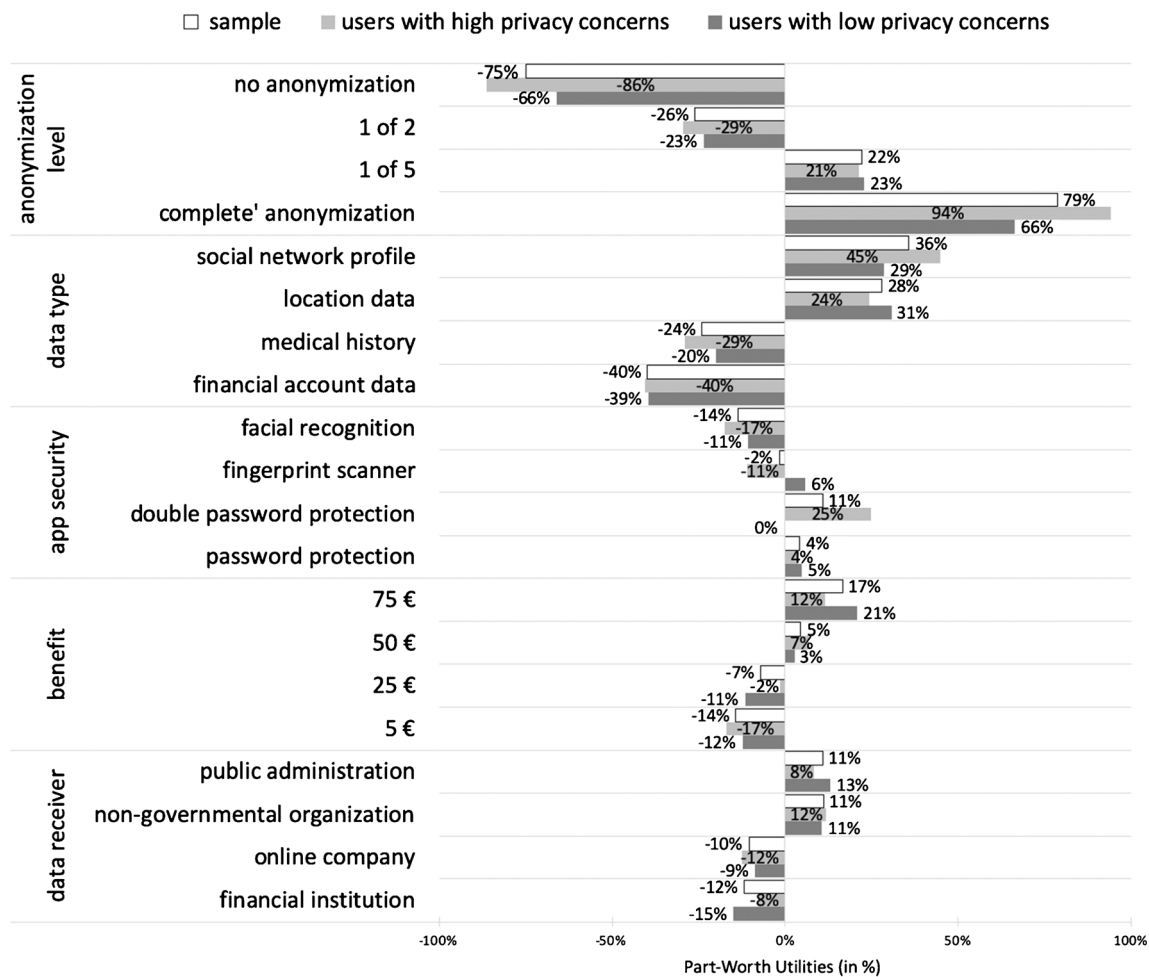


Fig. 10 Part-worth utilities of all attribute levels (n = 126)

topic of privacy in data-sharing which provides, on the one hand, a deeper understanding of users' notions of privacy in data sharing, and, on the other hand, a quantification of users' preferences and trade-offs in data sharing decisions.

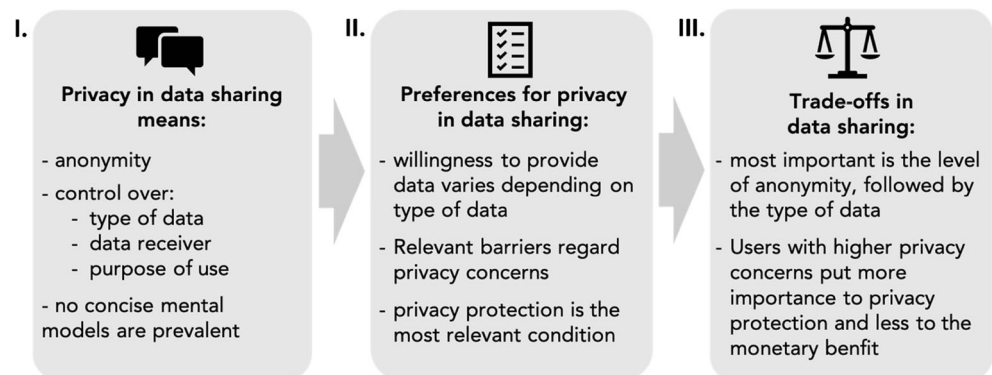
Identity and Anonymization

The findings of all three studies indicate that the protection of one's identity is the core element for privacy and,

correspondingly, anonymization is the most important factor that influences users' decisions to share data. The results are important for the design of privacy-preserving data markets as they show that good anonymization techniques are a core element to ensure users' perceived privacy.

In Study I, a qualitative approach was used to identify mental models of and preferences for privacy in data sharing. To reach the desired level of privacy – that is a balance between withdrawal and self-disclosure – we learn to use

Fig. 11 Key findings of the consecutive research process



physical mechanisms in the ‘offline’ world, like fences, doors, and curtains. In the online world and especially regarding data sharing contexts, these binary approaches are not fully applicable. Privacy is often misunderstood as the complete withdrawal and the so-called privacy paradox is observable when users decide to share data, e.g., on social media. But privacy is a state that is managed via withdrawal and self-disclosure. Self-disclosure is a precondition for social identity and relationships (Lahlou 2008; Taddicken 2014). Our results show that when giving users informational self-determination and anonymized sharing of data, they perceive their privacy to be guarded.

To quantify attitudes and visions about online privacy, the results and the importance of anonymization were confirmed in a subsequent survey study (Study II), including a large German sample ($N = 800$ participants). Anonymization is rated as an important condition for privacy in data sharing, but other factors were evaluated similarly important, especially data security, the ways and the possibility of deleting data, and transparency. However, evaluating single factors in isolation, as it is typically done in surveys, does not tell us how users would decide in real world settings. People need to decide under which circumstances they would be willing to share data and, also, under which conditions they would not do so. Thus, they weigh the positive and the negative aspects against each other. In order to study these trade-offs, a conjoint experiment (Study III) was developed in which the participants chose between scenarios of data sharing that differed in anonymization level, type of data, data receiver, amount of monetary rewards, and app security. This approach reveals the trade-offs between the factors and what impact each factor has on the decision to share data. The conjoint analysis demonstrated that anonymization level was the most decisive factor for the decision to share data. This result is in line with previous research on data sharing in a medical context (Calero Valdez and Ziefle 2019). Other previous results of Conjoint experiments which showed the importance of rewards and costs for data sharing (Krasnova et al. 2009; Roeber et al. 2015) decisions were not supported. These differences could be attributed to different samples and cultural settings as well as different contexts of data sharing. Also, the selection and operationalization of attributes and levels differed.

Still, the applied conjoint approach has some limitations. The number of attributes was already much for conjoint tasks, as the participants should not be overtaxed in their abilities to validly evaluate complex scenarios. A trade-off decision between the complexity of the research issue and an economic design was made, although Studies I and II showed that still more aspects are important to the participants, e.g., deletion of data, storage conditions, and control of the purpose of data collection. With the use of adaptive conjoint approaches, more attributes could be included into future studies. Additionally, the comparatively small and young sample of the conjoint

study needs to be taken into account. Empirical research has found that privacy perceptions and behaviors differ between age groups (Miltgen and Peyrat-Guillard 2014; Van den Broeck et al. 2015). Future studies should therefore include a wider span of age and focus also on the youngest and the oldest that have been described as vulnerable groups regarding online privacy (Moscardelli and Divine 2007; Van den Broeck et al. 2015).

The levels of the attribute anonymization were chosen to be comprehensible to participants without background knowledge of anonymization techniques. The level of ‘complete’ anonymization does not exist in reality and was only used to compare it to k -anonymity levels of $k = 5$ and $k = 2$. By that, we see that there is large gap in utility between an anonymization of “1 of 5” to “complete anonymization”, emphasizing that “1 of 5” is not perceived as anonymization level equal to complete anonymization. This small deception may still have influenced the evaluation of the other presented anonymization levels. Future studies should try to understand users’ risk perceptions in data sharing and how the perception of anonymization level is distributed.

Further important aspects for user-centered privacy-preserving data markets

The qualitative approach revealed many aspects that are important to users for such data markets. Informational self-determination – meaning in this context giving users full control over what data they share with whom for what purpose – and trust in the data market, its data security, and the handling of the data by the data receivers are the most important factors from the users’ perspective besides anonymization. The most relevant barriers against using a privacy-preserving data market are privacy concerns or mistrust into the data security of data markets. Moral concerns about the trade with personal data are also influential. Full transparency and control by the users are needed to build trust. Users want to know what their data are used for and thus observe the benefit besides the monetary reward.

The willingness to provide data varies between data types and the type of data is the second most important attribute of the conjoint study. The sensitivity of information plays an important role for the risk perceptions and the willingness to share data (Calero Valdez and Ziefle 2019; Milne et al. 2016; Wirth et al. 2019). Users are willing to provide (certain) data when it has a benefit for themselves or the society and when trust in the data protection is prevalent. These trade-offs between privacy and utility goes in line with the privacy calculus theory (Dinev and Hart 2006) and can thus explain the ‘paradoxical’ behavior of internet users providing data even if they are concerned about their privacy. The operationalization of benefit in this study as monetary reward between 5€ and 75€ may also have influenced the importance scores. If the type of

benefit would have included rewards of 1000€, we would hypothesize higher importance scores for monetary reward. Also, the value of money is different for each participant and some participants may wish for other benefits than money. All in all, the findings of this study only apply to the investigated attributes and levels.

But the results show that privacy is very complex. The new European Data Protection Regulation (GDPR) applies only to personal data which is defined as any information relating to an identified or identifiable natural person (cf. Art. 4 (1) GDPR). Thus, for the use of anonymized data, no consent is needed by the individual. The results of our study show that not only anonymization is important for informational self-determination. Even with anonymization, users want to control what data about them are used for which purposes and they want to be rewarded. Here, privacy-preserving data markets can offer a solution to restore the users' informational self-determination and let users get a share of the huge profits made with their data.

In contrast to data monopolies by large online companies who operate as 'walled gardens', making anonymized, high quality user data accessible with a low market entry barrier, higher legal certainty for the buyers, and transparency for all involved parties offers great chances to companies and lays the foundation for innovations. But there are challenges for privacy-preserving data markets. For one, anonymized data may lose value to paying companies and thus the data may represent a bad alternative compared to existing data repositories. However, customers might prefer companies accepting their boundaries and need for privacy, thus making the use of a privacy-preserving data market as source for marketing data a relevant business advantage. For another, to offer a viable and competitive data repository, a minimum amount of user data needs to be accumulated, thereby making the initiation phase critical for such a privacy-preserving data market. This argument raises the question whether a privacy-preserving data market – which offers benefits and informational self-determination for users as well as innovation potential for small and medium sized companies – could or should be provided by a governmental or a non-profit organization. Initiatives into a similar direction have been started in Europe (Poikola et al. 2015; Matzutt et al. 2017). Our results show that users are willing to share data on their terms when provided with informational self-determination and, particularly, anonymization. This holds especially true for those users with less privacy concerns.

Still, it has to be considered that the results of our studies show preference ratings and not actual behavior. In privacy research and other contexts, the discrepancy between stated attitudes and actual behavior has often been revealed (Gerber et al. 2018). Thus, the findings of this study indicate that anonymization level is very important to the participants, the medical history is shared less willingly than location data, and public institutions are more

accepted as data receivers than companies. But actual behavior might present differently according to the privacy calculus theory. This may implicate that attractive benefits may override privacy concerns. Also, the privacy calculus is influenced by situational and affective variables (Acquisti et al. 2015; Kehr et al. 2015). During our empirical research, regardless whether it is within a focus group, online questionnaire, or conjoint study, the participants are primed to privacy as topic and answered mostly rationally. This may not be the case in real data sharing decisions, where the immediate benefit is more foreground. Future studies therefore need to explore how users might be adequately warned or informed about their privacy settings. Here, user-centered recommendation systems might prevent imprudent online behaviors by illustrating consequences for possible privacy intrusions.

User diversity

Throughout all three studies, differences between groups of participants with lower and higher privacy concerns could be observed. Users with lower privacy concerns in general are more motivated to use privacy-preserving data markets and less concerned for their privacy when doing so. They are probably the first user group to participate in such data markets. This shows how diverse users are in their privacy attitudes and preferences and that different user groups exist and need to be considered in research and implementation (e.g., Schomakers et al. 2019a; Sheehan 2002). Here, only these two groups were studied, but users' preferences probably vary also depending on other characteristics and attitudes, e.g., experiences, knowledge, or trust (e.g., Riquelme and Román 2014).

Also, the cultural setting needs to be kept in mind: All three studies were conducted in Germany. Historical experiences with dictatorship and surveillance as well as a long tradition of privacy regulation may have an influence on Germans' privacy attitudes. Empirical studies have shown that Germans perceive more risks in comparison to other countries and perceive information to be more sensitive (Krasnova and Veltri 2010; Trepte et al. 2017; Schomakers et al. 2019b). Comparison of results to other countries would be insightful.

Acknowledgements We thank all participants of the studies for their openness to share their thoughts and opinions about privacy and its protection. Parts of this work have been funded by the German Ministry of Education and Research (BMBF) under project MyneData (KIS1DSD045).

Funding Information Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included

in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behaviour in the age of information. *Science*, 347(6221), 509–514. <https://doi.org/10.2139/ssm.2580411>.
- Altman, I. (1976). Privacy - a conceptual analysis. *Environment and Behavior*, 8(1), 7–29.
- Arning, K. (2017). Conjoint measurement. *The International Encyclopedia of Communication Research Methods*, 1–10. <https://doi.org/10.1002/9781118901731.iecrm0040>.
- Asgharpour, F., Liu, D., & Camp, L. J. (2007). Mental Models of Security Risks. In S. Dietrich & R. Dhamija (Eds.), *Financial Cryptography and Data Security: 11th International Conference, FC 2007, and 1st International Workshop on Usable Security, USEC 2007, Scarborough, Trinidad and Tobago, February 12–16, 2007. Revised Selected Papers* (pp. 367–377). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77366-5_34.
- Baruh, L., & Popescu, M. (2017). Big data analytics and the limits of privacy self-management. *New Media & Society*, 19(4), 579–596. <https://doi.org/10.1177/1461444815614001>.
- Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication*, 67, 26–53. <https://doi.org/10.1111/jcom.12276>.
- Blank, G., Bolsover, G., & Dubois, E. (2014). *A new privacy paradox: Young people and privacy on social network sites*.
- Calero Valdez, A., & Ziefle, M. (2019). The users' perspective on the privacy-utility trade-offs in health recommender systems. *International Journal of Human-Computer Studies*, 121, 108–121. <https://doi.org/10.1016/j.ijhcs.2018.04.003>.
- Camp, L. J. (2009). Mental models of privacy and security. *IEEE Technology and Society Magazine*, 28(3), 37–46. <https://doi.org/10.1109/MTS.2009.934142>.
- Conger, S., Pratt, J. H., & Loch, K. D. (2013). Personal information privacy and emerging technologies. *Information Systems Journal*, 23(5), 401–417. <https://doi.org/10.1111/j.1365-2575.2012.00402.x>.
- Coopamootoo, K. P. L., & Groß, T. (2014). Mental models for usable privacy: A position paper. In T. Tryfonas & I. Askoxylakis (Eds.), *Human aspects of information security, privacy, and trust: Second international conference, HAS 2014, held as part of HCI international 2014, Heraklion, Crete, Greece, June 22–27, 2014. Proceedings* (pp. 410–421). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-07620-1_36.
- 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>.
- Dourish, P., Delgado De La Flor, J., & Joseph, M. (2003). Security as a Practical Problem: Some Preliminary Observations of Everyday Mental Models. *Proceedings of CHI 2003 Workshop on HCI and Security Systems*.
- Egelman, S., Felt, A. P., & Wagner, D. (2013). Choice architecture and smartphone privacy: There's a price for that. In *The economics of information security and privacy* (pp. 211–236). Springer, Berlin, Heidelberg.
- Enserink, M., & Chin, G. (2018). The end of the privacy. *Science*, 347(6221), 490–491. <https://doi.org/10.4324/9780203119815-5>.
- European Commission. (2015). Data protection Eurobarometer. Retrieved from <https://www.gov.uk/data-protection/the-data-protection-act>
- Gentner, D., & Stevens, A. L. (Eds.). (2014). *Mental models*. Psychology Press.
- Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77, 226–261. <https://doi.org/10.1016/j.cose.2018.04.002>.
- Gkatzelis, V., Aperjis, C., & Huberman, B. A. (2015). Pricing private data. *Electronic Markets*, 25(2), 109–123. <https://doi.org/10.1007/s12525-015-0188-8>.
- Gustafsson, A., Herrmann, A., & Huber, F. (Eds.). (2007). *Conjoint measurement: methods and applications*. Springer Science & Business Media.
- Hallam, C., & Zanella, G. (2017). Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. *Computers in Human Behavior*, 68, 217–227. <https://doi.org/10.1016/j.chb.2016.11.033>.
- 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), 1–18. <https://doi.org/10.5817/CP2016-4-7>.
- Hornung, G., & Schnabel, C. (2009). Data protection in Germany I: The population census decision and the right to informational self-determination. *Computer Law and Security Review*, 25(1), 84–88. <https://doi.org/10.1016/j.clsr.2008.11.002>.
- Hsieh, H.-F., & Shannan, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>.
- 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). <https://doi.org/10.1111/isj.12062>.
- Kokolakis, S. (2015). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 2011(2013), 1–29. <https://doi.org/10.1016/j.cose.2015.07.002>.
- Krasnova, H., & Veltri, N. F. (2010). Privacy Calculus on social networking sites: Explorative evidence from Germany and USA. *Proceedings of the annual Hawaii international conference on system sciences*, 1–10. <https://doi.org/10.1109/HICSS.2010.307>.
- Krasnova, H., Hildebrand, T., & Guenther, O. (2009). *Investigating the Value of Privacy in Online Social Networks: Conjoint Analysis*. ICIS 2009 Proceedings. Paper 173. <http://aisel.aisnet.org/icis2009/173>. Accessed 4 Feb 2020.
- Lahlou, S. (2008). Identity, social status, privacy and face-keeping in digital society. *Social Science Information*, 47(3), 299–330. <https://doi.org/10.1177/0539018408092575>.
- Li, Y. (2011). Empirical studies on online information privacy concerns: Literature review and an integrative framework. *Communications of the Association for Information Systems*, 28(28), 453–496.
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1(1), 1–27. [https://doi.org/10.1016/0022-2496\(64\)90015-X](https://doi.org/10.1016/0022-2496(64)90015-X).
- Lutz, C., & Strathoff, P. (2013). Privacy concerns and online behavior – Not so paradoxical after all? *Multinationale Unternehmen Und Institutionen Im Wandel – Herausforderungen Für Wirtschaft, Recht Und Gesellschaft*, 81–99. <https://doi.org/10.2139/ssrn.2425132>.
- Markos, E., Milne, G. R., & Peltier, J. W. (2017). Information sensitivity and willingness to provide continua: A comparative privacy study of

- the United States and Brazil. *Journal of Public Policy & Marketing*, 36(1), 79–96. <https://doi.org/10.1509/jppm.15.159>.
- Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114–133. <https://doi.org/10.1177/1461444810365313>.
- Matzutt, R., Müllmann, D., Zeissig, E. M., Horst, C., Kasugai, K., Lidynia, S., ... & Ziefle, M. (2017). myneData: towards a trusted and usercontrolled ecosystem for sharing personal data. Maximilian Eibl, Martin Gaedke (Eds.): INFORMATIK 2017, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik.
- McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, 13(3), 334–359. <https://doi.org/10.1287/isre.13.3.334.81>.
- Milne, G. R., Pettinico, G., Hajjat, F. M., & Markos, E. (2016). Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs*, 1–29. <https://doi.org/10.1111/joca.12111>.
- Miltgen, C. L., & Peyrat-Guillard, D. (2014). Cultural and generational influences on privacy concerns: A qualitative study in seven European countries. *European Journal of Information Systems*, 2317(2), 103–125. <https://doi.org/10.1057/ejis.2013.17>.
- Morgan, M. G., Fischhoff, B., Bostrom, A., & Atman, C. J. (2002). *Risk communication: A mental models approach*. Cambridge: Cambridge University Press.
- Moscardelli, D. M., & Divine, R. (2007). Adolescents' concern for privacy when using the internet: An empirical analysis of predictors and relationships with privacy-protecting behaviors. *Family and Consumer Sciences Research Journal*, 35(3), 232–252. <https://doi.org/10.1177/1077727X06296622>.
- 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>.
- Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford: Stanford University Press. https://doi.org/10.1207/S15327051HC116234_03.
- Palen, L., & Dourish, P. (2003). Unpacking “privacy” for a networked world. *Proceedings of the Conference on Human Factors in Computing Systems - CHI '03*, (5), 129. <https://doi.org/10.1145/642633.642635>.
- Poikola, A., Kuikkaniemi, K., & Honko, H. (2015). *Mydata a nordic model for human-centered personal data management and processing*. Finnish Ministry of Transport and Communications.
- Prettyman, S. S., Furman, S., Theofanos, M., & Stanton, B. (2015). Privacy and security in the brave new world: The use of multiple mental models. In: *International Conference on Human Aspects of Information Security, Privacy, and Trust* (pp. 260–270). Springer, Cham.
- Riquelme, I. P., & Román, S. (2014). Is the influence of privacy and security on online trust the same for all type of consumers. *Electronic Markets*, 24(2), 135–149. <https://doi.org/10.1007/s12525-013-0145-3>.
- Roeber, B., Rehse, O., Knorrek, R., & Thomsen, B. (2015). Personal data: How context shapes consumers' data sharing with organizations from various sectors. *Electronic Markets*, 25(2), 95–108. <https://doi.org/10.1007/s12525-015-0183-0>.
- Schomakers, E-M., Lidynia, C. & Ziefle, M. (2018). Hidden within a Group of People - Mental Models of Privacy Protection. In *Proceedings of the 3rd International Conference on Internet of Things, Big Data and Security (IoTBDS 2018)*, 85–94. <https://doi.org/10.5220/0006678700850094>.
- Schomakers, E-M., Lidynia, C., & Ziefle, M. (2019a) A Typology of Online Privacy Personalities. *Journal of Grid Computing* 17(4): 727–747.
- Schomakers, E-M., Lidynia, C., Müllmann, D., & Ziefle, M. (2019b) Internet users' perceptions of information sensitivity – insights from Germany. *International Journal of Information Management* 46: 142–150.
- Sheehan, K. B. (2002). Toward a typology of internet users and online privacy concerns. *The Information Society*, 18(1), 21–32. <https://doi.org/10.1080/01972240252818207>.
- Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4), 989–1015.
- Spiekermann, S., & Novotny, A. (2015). A vision for global privacy bridges: Technical and legal measures for international data markets. *Computer Law and Security Review*, 31(2), 181–200. <https://doi.org/10.1016/j.clsr.2015.01.009>.
- Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K. L. (2015). The challenges of personal data markets and privacy. *Electronic Markets*, 25(2), 161–167. <https://doi.org/10.1007/s12525-015-0191-0>.
- 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-Disclosure1. *Journal of Computer-Mediated Communication*, 19(2), 248–273. <https://doi.org/10.1111/jcc4.12052>.
- Trepte, S., Reinecke, L., Ellison, N. B., Quiring, O., Yao, M. Z., & Ziegele, M. (2017). A cross-cultural perspective on the privacy Calculus. *Social Media + Society*, 3(1). <https://doi.org/10.1177/2056305116688035>.
- Van den Broeck, E., Poels, K., & Walrave, M. (2015). Older and wiser? Facebook use, privacy concern, and privacy protection in the life stages of emerging, young, and middle adulthood. *Social Media + Society*, 1(2). <https://doi.org/10.1177/2056305115616149>.
- Westin, A. F. (1967). Privacy and Freedom. *American Sociological Review*, 33(1), 173. <https://doi.org/10.2307/2092293>.
- Wirth, J., Maier, C., Laumer, S., & Weitzel, T. (2019). Perceived information sensitivity and interdependent privacy protection: A quantitative study. *Electronic Markets*, 29, 359–378. <https://doi.org/10.1007/s12525-019-00335-0>.
- Xu, H., Dinev, T., Smith, H.J., & Hart, P. (2008). "Examining the Formation of Individual's Privacy Concerns: Toward an Integrative View". ICIS 2008 Proceedings. Paper 6. <http://aisel.aisnet.org/icis2008/6>. Accessed 4 Feb 2020.
- Ziefle, M., Halbey, J., & Kowalewski, S. (2016). Users' willingness to share data in the Internet: Perceived benefits and caveats. In *Proceedings of the International Conference on Internet of Things and Big Data (IoTBD 2016)* 255–265.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.