Research Article

Beware: Processing of Personal Data—Informed Consent Through Risk Communication

—Lukas Seiling, Rita Gsenger, Filmona Mulugeta, Marte Henningsen, Lena Mischau, and Marie Schirmbeck

***Abstract*—Background:** *The General Data Protection Regulation (GDPR) has been applicable since May 2018 and aims to further harmonize data protection law in the European Union. Processing personal data based on individuals' consent is lawful under the GDPR only if such consent meets certain requirements and is "informed," in particular. However, complex privacy notice design and individual cognitive limitations challenge data subjects' ability to make elaborate consent decisions. Risk-based communication may address these issues.* **Literature review:** *Most research focuses on isolated aspects of risk in processing personal data, such as the actors involved, specific events leading to risk formation, or distinctive (context-dependent) consequences. We propose a model combining these approaches as the basis for context-independent risk communication.* **Research questions:** *1. What are relevant information categories for risk communication in the processing of personal data online? 2. Which potentially adverse consequences can arise from specific events in the processing of personal data online? 3. How can consequences in the processing of personal data be avoided or mitigated?* **Research methodology:** *The GDPR was examined through a systematic qualitative content analysis. The results inform the analysis of 32 interviews with privacy, data protection, and information security experts from academia, Non-Governmental Organizations, the public, and the private sector.* **Results:** *Risk-relevant information categories, specific consequences, and relations between them are identified, along with strategies for risk mitigation. The study concludes with a specified framework for perceived risk in processing personal data.* **Conclusion:** *The results provide controllers, regulatory bodies, data subjects, and experts in the field of professional communication with information on risk formation in personal data processing. Based on our analysis, we propose information categories for risk communication, which expand the current regulatory information requirements.*

Index Terms*—Data protection, general data protection regulation (GDPR), informed consent, privacy notice, risk communication, risk model.*

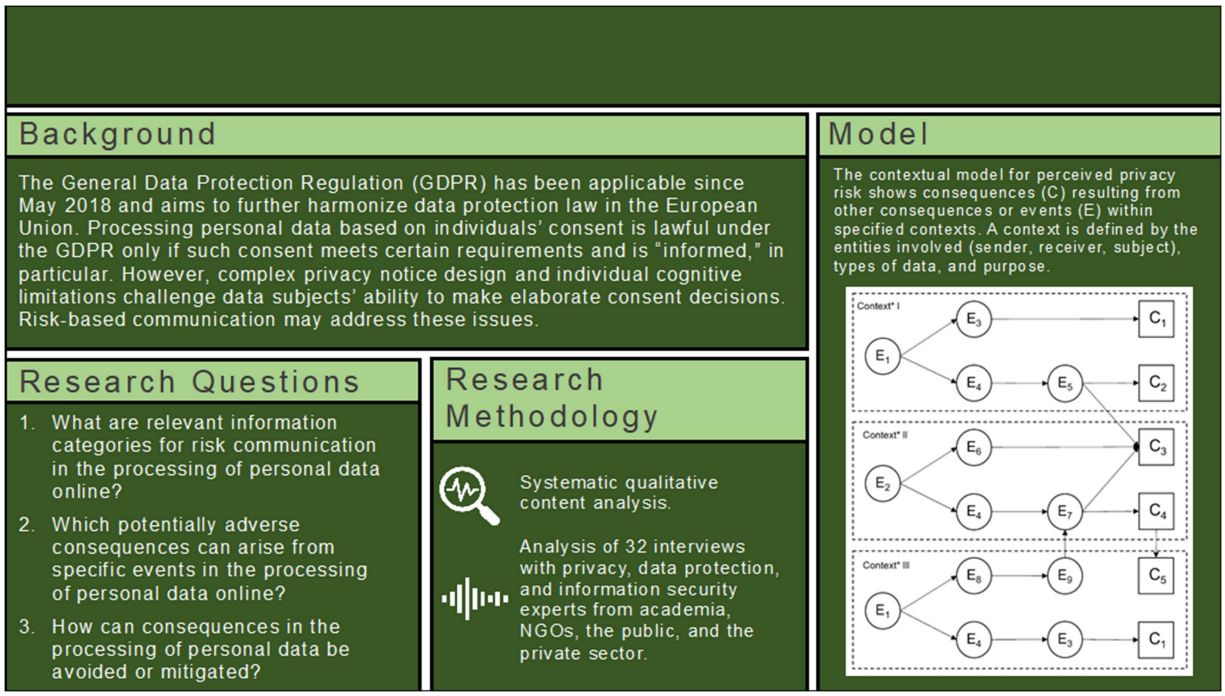
Background

This section introduces the General Data Protection Regulation (GDPR), which provides the fundamental legal framework for processing personal data in the European Union. It briefly summarizes known issues with informed and rational decision making based on privacy policies, and further reviews previous attempts to address existing limitations through communicative means. Finally, it proposes risk communication as a promising additional measure for practitioners to ensure privacy awareness and better-informed consent decisions.

Manuscript received 12 April 2023; revised 24 October 2023; accepted 25 October 2023. Date of publication 13 March 2024; date of current version 14 March 2024. *(Corresponding author: Lukas Seiling.)*
Lukas Seiling and Rita Gsenger are with the Weizenbaum Institute for the Networked Society, 10623 Berlin, Germany (email: lukaseiling@weizenbaum-institut.de; rita.gsenger@weizenbaum-institut.de).
Filmona Mulugeta is with Mercedes-Benz Tech Innovation, 10243 Berlin, Germany (email: filmona.mulugeta@mercedes-benz.com).
Marte Henningsen is with the Institut für Kognitionswissenschaft, 49090 Osnabrück, Germany (email: mhenningsen@uni-osnabrueck.de).
Lena Mischau and Marie Schirmbeck were with the Weizenbaum Institute for the Networked Society, 10623 Berlin, Germany (email: lena.mischau@web.de; marie.schirmbeck@gmail.com).
This paper has supplementary downloadable material at https://ieeexplore.ieee.org provided by the authors. The PDF file consists of Appendixes A-C (332 kB in size).

IEEE 10.1109/TPC.2024.3361328

GDPR and Informed Consent The protection of natural people in relation to the processing of personal data is a fundamental right in the European Union (Rec 1) [1]. To further harmonize data protection law, the European Union adopted the GDPR. The GDPR has been directly applied to all European Union member states since May 25, 2018, and aims to increase legal certainty to ensure a consistent and high level of data protection and to remove the obstacles to flows of personal data within the European Union [Art. 99 (2), Rec. 10] [1]. To meet these objectives, the GDPR establishes a detailed legal framework that includes, among others, specific data protection principles (Art. 5) and various obligations for the controller, i.e., the natural or legal person, public authority, agency, or other body, which determines the purposes and



means of the processing of personal data in each case [Art. 4 (7)]. Personal data must be processed lawfully, fairly, and in a transparent manner, in particular [1, Art. 5 (1)] [1], [3].

This means that personal data may only be processed if certain requirements are met. Most important, personal data may only be processed if there is a legal ground for processing pursuant to the GDPR (Art. 6 and Art. 9) and if the individual, whose personal data are processed—the so-called data subject [Art. 4 (1)] is informed about the details of such processing (Art. 12–14). Processing may be lawful, for example, if the data subject has given consent to such processing of his or her personal data for one or more specific purposes [Art. 6 (1)(a)]. If processing is based on consent, the GDPR requires that consent must be, among others, freely given and informed [Art. 4 (11), Art. 7(2), Rec. 32] [1].

Limitations of Informed Consent Data controllers implement the GDPR's provisions concerning informed consent under the notice and choice paradigm [4]. It assumes that the privacy notice providing all relevant information about the processing of personal data enables the data subject to make a rational and informed decision. However, the design of privacy notices suffers from widely documented malpractice [5], [6]. Disregarding such negative examples, users and experts often still have issues understanding these documents [3]. Various design issues [7], [8] contribute to this confusion and result in most privacy notices being left unread [9].

Nevertheless, even if privacy notices are read and understood, people have limitations of cognitive capacity that influence their rational decision making [10], [11]. A manifestation of these limitations is, for example, "hyperbolic discounting" [12, p. 106], the overestimation of immediate consequences and discount of future consequences [13], [14]. These limitations complicate the balanced benefit and risk assessment necessary for accurate decision-making processes [15], [16].

Although potential benefits are most evident from a service's purpose, privacy risk information is not always available and accessible [17], [15]. Empirical evidence indicates that people lack knowledge of potential privacy consequences because of inaccurate or incomplete mental models about the controllers' data processing [18], [19]. Many data subjects do not consider the consequences of granting or refusing consent. Therefore, professional communication is necessary to support less biased consent decisions and increase privacy awareness [20].

Informed Consent through Risk Communication Various projects have tried to design evaluative measures to increase privacy awareness, such as communication akin to nutrition labels [22], [23] or Privacy Bird, an audiovisual warning system using bird imagery to indicate whether a website's privacy

Practitioner Takeaway

- Risk communication can address user trust by providing evaluations and justifications for privacy policies.
- Data processing operations are the starting point of risk formation and should therefore be central to risk communication.
- Although consequences might be difficult to communicate, the Contextual Model for Perceived Privacy Risk addresses this issue by linking latent consequences with more explicit, tangible consequences, and highlighting means of mitigating negative consequences.

policy matches user preferences [21]. Another approach is the use of privacy icons "to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing" (Art. 12 (7), GDPR) [1]. Researchers [24], [25], nongovernmental organizations, and citizen initiatives [26], [27], [28] proposed multiple icon sets based on the regulation.

Topic icons have dominated research and publications on privacy icons in the last decade. Those serve as

> information markers that effectively support the navigation through large amounts of legal information and increase speed and accuracy of comprehension. [25, p. 193]

However, they communicate only headline information and lack any evaluation of consequences and risks, therefore holding limited informational value. Furthermore, when applied to privacy policies, topic icons still require data subjects to at least partially read privacy notices to seek out the controller's practices relevant to them.

The GDPR [1] states that data subjects "should be made aware of risks … in relation to the processing of personal data" (Rec. 39) and to obliges controllers to inform data subjects about the details of the way their personal data are processed Art. 12 to 14. We therefore suggest examining how risk communication may support data subjects in their privacy choices. Therefore, this article adopts a risk-based approach [29] to identify relevant information categories for risk communication.

At first glance, this approach seems likely to be opposed by data controllers [30]. However, risk communication does not necessarily lessen the likelihood of using a service. Instead, users might also interpret it as a reflecting concern for their privacy [31]. Effective risk communication can therefore help controllers and professional communication experts increase their trustworthiness, especially if there are few risks associated with the use of a service [48].

Literature Review

This literature review first presents different risk conceptualizations while establishing the foundational definition for the rest of this article. Afterward, the Perceived Risk Model is introduced to understand how people perceive (privacy) risks. In the subsequent section, the model is expanded with the concept of Contextual Privacy, which allows integrating different contexts into the model. By combining these approaches, we develop the Contextual Model for Perceived Privacy Risk with three critical dimensions of analysis: events, consequences, and contexts. The last section reviews the warning literature to identify additional requirements for successful risk communication.

Defining Privacy Risk Agreeing upon a definition of risk in personal data processing is not trivial. The concept of privacy, and thereby privacy risk, is "essentially contested" [32]. Furthermore, this process is complicated by multifaceted approaches to privacy risk, which take legal, economic, societal, or software engineering perspectives [49].

Most approaches to privacy risk share a common understanding of risk, as captured by the International Organization for Standardization, which defines risk as

> the effect of uncertainty on objectives … often characterized by reference to potential events and consequences … expressed in terms of a combination of the consequences of an event … and the associated likelihood of occurrence. [33]

However, the numeric expression of risk regarding privacy based on information about the likelihood and severity of privacy-relevant consequences [34] is challenging.

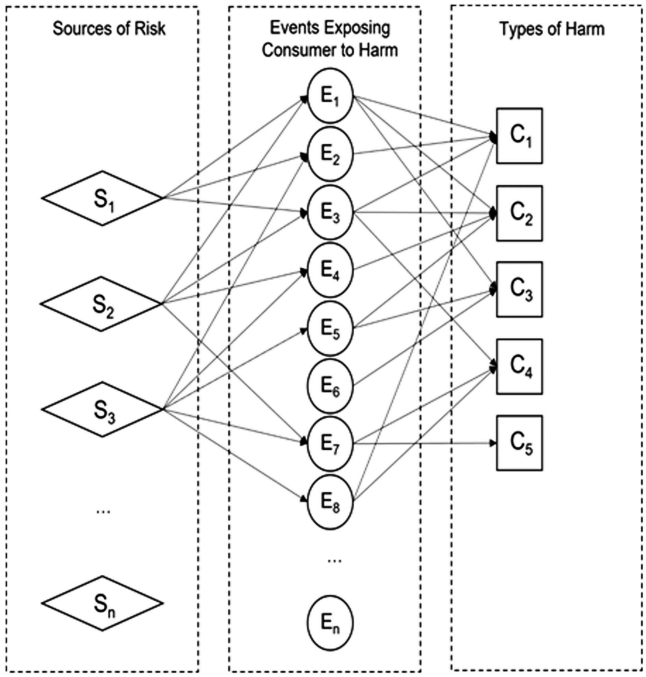


Fig. 1. Perceived Risk Model as proposed by Glover and Benbasat [36].

Severity evaluations may differ between individuals, and the lack of a standardized set of consequences prevents an empirical likelihood estimation [35]. Therefore, this article pursues a less formalized approach, focusing on events ("occurrence[s] or change[s] of a particular set of circumstances" [33]) and consequences ("outcome[s] of an event" [33]) in the processing of personal data to capture relevant information categories for the communication of privacy risks.

Conceptualizing Perceived Privacy Risk Glover and Benbasat [36] differentiated perceived risk from actual risk. Since people "cannot respond to what they do not perceive," they propose to study "a person's perception of the uncertain and adverse consequences of engaging in an activity" [36, p. 48]. Understanding these perceived risks is especially important for risk communication, as risk perception impacts privacy intentions [37], [38] and behavior [39], [40].

The *Model of Perceived Risk* [36] provides a means to understand risk in the processing of personal data by describing how a data subject could experience harm from a transaction (as shown in Fig. 1).

1. Some phenomenon or actor is the *source* of the risks.
2. The data subject will suffer harm from that source only if an *event* exposing the data subject to harm occurs.
3. An event may result in one or more *types of harm* to the data subject.

According to Glover and Benbasat's Perceived Risk Model, potential sources of risks can be actors, such as retailers, product manufacturers, or unknown third parties, which, through specific events, can expose consumers to different kinds of harm on a financial, temporal, psychological, social, or physical dimension [36].

To illustrate the model, consider the following example: A social network provider (the source) might have to provide enforcement services with personal message data of a data subject (the event), which might result in criminal prosecution and punishment (types of harm) for the data subject. For instance, Celeste Burgess, a teenager from Nebraska, was sentenced to 90 days in jail for violating federal abortion laws after police read her private Facebook messages [41].

Glover and Benbasat [36] stressed that it is crucial to collectively study sources, events, and resulting consequences collectively to provide data subjects with complete of risk and enable risk perception.

Contextualizing Perceived Privacy Risk The model of perceived risk does not account for contextual differences. However, privacy is greatly contextual, exemplified by the specific challenges in processing personal data for research or health-related purposes [43]. To account for contextual differences in perceived privacy risk, we expand the model of perceived risk by drawing on Nissenbaum's concept of privacy as contextual integrity [42], [44]. A context describes a structured social setting contingent on time, place, culture, and other aspects. It can be defined through its ends and values (i.e., its purposes), the different entities involved in information exchange (the sender, the receiver, and the subject whom the information is about), and the type of information exchanged. What exchanges are considered appropriate depends on the privacy norms of the context [44].

This expansion yields the contextual model for perceived privacy risk (see Fig. 2). It contains events that might violate privacy norms and lead to harm, or, more generally, negative or unexpected consequences. All events occur within a context specified by the entities involved, the processing

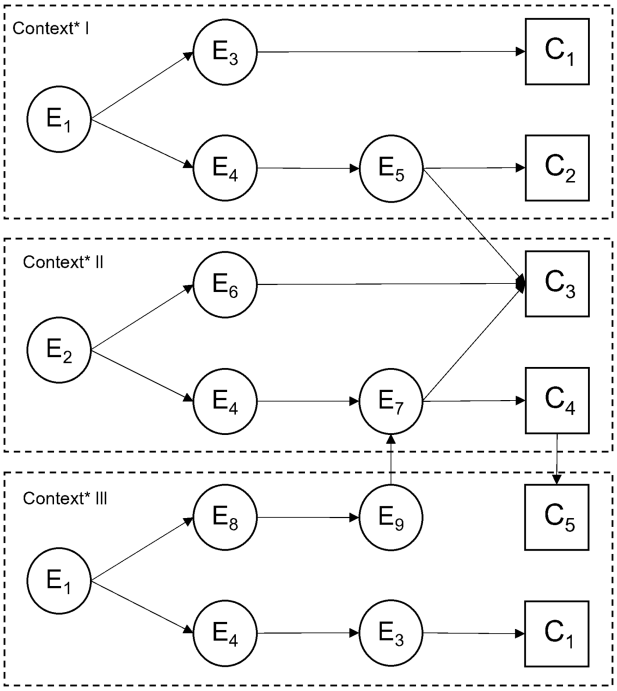


Fig. 2. Contextual model for perceived privacy risk (original figure for this publication). *The context is defined by the entities involved (sender, receiver, and subject), types of data, and purpose.

purposes, and the types of personal data processed. Thus, the actor as the source of adverse events is not a separate entity but a necessary aspect of the context within which data are processed.

That model allows for a new set of assumptions.

1. Events can lead to the *transgression of contexts*, indicated by the arrows between event E5 and consequence C3 within a different context.
2. Consequences can have cascading effects, causing *second-order consequences*, indicated by the arrow between consequences C4 and C5.
3. The *same consequence* can manifest itself in *multiple contexts*, indicated by the consequence C1 appearing in both Contexts I and III.

To illustrate these assumptions—and notwithstanding desired advantages of algorithmically supported decisions—consider the case of Robert Julian–Borchak Williams, a Black man who was falsely identified by facial recognition software [45] used by the police department, which generated a match between surveillance footage and the portrait on Williams' driver's license. Using the picture on the driver's license by the police in such a manner constitutes a transgression of contexts. There is a change in actors (administration → police), purposes (authorization for operating a motorized vehicle → criminal investigation), and types of data (data relating to driving authorization → data relating to criminal offenses). The generated match was false (emergence of defective data as a first-order consequence) and led to a 30-hour detention (stigmatization as a second-order consequence). Although the specific stigmatization occurred within the context of criminal prosecution, it could also have led to stigmatization in a work context or other social contexts (thereby potentially manifesting in multiple contexts).

Following this model, research on risk in processing personal data should account for events, consequences, and relations between them, while also considering similarities and differences across contexts. However, few publications focus on the entities involved or specific events [18], [46] and most address various (context-dependent) consequences that result from data processing [39], [46]. Only a small subset of empirical studies explicitly link events to consequences in processing of personal data [46], [47].

Communicating Privacy Risk Risk is usually communicated through warnings. The Communication–Human Information Processing (C-HIP) Model [48] proposes a two-part framework to describe warning processes: the first part describes the communication of a warning message (including the source and the channel, or media, through which the warning is transmitted, as well as its content). The second part focuses on how warnings are processed by a receiver (influenced by attention, comprehension, attitudes, and motivation, possibly resulting in behavioral compliance) [48]. The rest of this article focuses exclusively on the first section of this model, more specifically, message content.

A warning should educate receivers about hazards (the specific circumstances that can result in negative or unexpected consequences) and appropriate behaviors. In addition, they should enable behavior modification to decrease incidents that result in adverse outcomes [49], [50]. Indeed, users are more interested in protecting themselves if they feel at risk and are aware of security mechanisms [51]. The risk, however, needs to be adequately communicated [51]. Consequence information is often not explicit and lacks necessary details [48]. Although people tend to be more aware of abstract privacy risks [52], they often cannot conclude specific risks or behavioral

consequences from more abstract risk information [46]. On the other hand, factual consequence information increases perceived hazards and cautious intent [53], as well as perceived severity.

Because few scientific works systematically address risk in processing personal data, we aim to identify risk-relevant information categories that should be included in risk communication aimed at data subjects. Crucially, this is not meant to replace privacy notices but to propose specific information categories that should be highlighted or prioritized for (supplementary) risk communication.

Research Questions

Given the seemingly infinite possibilities of data processing, we think it is useful to understand potential models of perceived risk and to communicate context-dependent and context-independent risks. We apply the *Contextual Model for Perceived Privacy Risk*, which specifies potential risk as relations between events and consequences in specific contexts to determine information categories required for adequate warnings (as proposed by the C-HIP model). According to Wogalter, a

> warning message should include information about the hazard, instructions on how to avoid the hazard, and the potential consequences if the hazard is not avoided. [48, p. 56]

Thus, a holistic approach to risk in the processing of personal data should strive to answer the following research questions.

> **RQ1.** What are relevant information categories for risk communication in the processing of personal data online?
>
> **RQ2.** Which potentially adverse consequences can arise from specific events in the processing of personal data online?
>
> **RQ3.** How can consequences in the processing of personal data be avoided or mitigated?

The most common methods to elicit information on perceived events or consequences are user focus groups and interviews [18], [36], [39], [46], [54]. Expert evaluations are rarely used [55]. To provide a conceptual basis for RQ1, we conducted a content analysis of the GDPR. To answer RQ2 and RQ3, we conducted and qualitatively analyzed expert interviews to identify additional categories, which provided a more nuanced understanding of the existing categories, potentially adverse consequences, and possible mitigation strategies. The methodology and results for the content analysis and the interviews are detailed below.

Research Methodology

This section explains the methodology of the studies in more detail. First, we describe the qualitative content analysis. Next, we present the expert interview methodology. This section includes further information about the recruitment and demographics of the interview participants, the interview procedure, and the data analysis.

Content Analysis To identify relevant information categories in the processing of personal data, we analyzed the GDPR using a systematic qualitative content (SQC) analysis based on Mayring [56]. A tentative coding agenda containing category definitions, anchor examples, and coding rules was designed based on previous privacy icons projects and criteria for high-risk processing operations [46]. Subsequently, a deductive and inductive analysis of 30% of the GDPR was conducted by a legal expert to provide a first set of categories. After that, three researchers with expertise in law, psychology, and engineering expanded the coding agenda by recoding the material. Finally, two other legal experts analyzed the complete GDPR according to the reworked coding agenda, including a sanity check after coding 30% of the material to prevent misunderstandings (for Cohen's Kappa, see Appendix A in the supplementary material, Tables VI and VII). The analysis resulted in a taxonomy that includes the following categories: data subject, data type, data processing, purpose of the data processing, and security (see Appendix A in the supplementary material, Table VIII).

Expert Interviews The content analysis results were validated by conducting expert interviews. Categories especially relevant for risk communication were determined by applying the Contextual Model of Perceived Privacy Risk to the transcripts. Furthermore, the qualitative interviews allowed for a more specific and contextualized assessment of the risk categories and their roles in risk communication.

Interview Participants: The expert interview participants were invited based on a previously held workshop [57], authorship of relevant literature, representation of privacy organizations, and the Slack instance of the MyData Conference

TABLE I
INTERVIEW PARTICIPANT DEMOGRAPHICS

	Interview Participants (n = 32)
Gender	
Male	19 (59%)
Female	11 (35%)
NA	2 (6%)
Discipline	
Interdisciplinary	13 (41%)
Law	18 (56%)
Social Sciences	10 (31%)
Comp. Science or Engineering	10 (31%)
Psychology	3 (9%)
Economics	3 (9%)
NA	1 (3%)
Area of employment	
Academia	18 (56%)
NGO	6 (19%)
Private Sector	6 (19%)
Public Sector	2 (6%)
Mean Years of Experience	9.59 (7.23)

Note: 41% of experts had an interdisciplinary background, which is why the disciplines add up to more than 100%.

2020. Experts were chosen as interview participants due to their knowledge about data protection and applicable legal regulations not accessible to laypeople [58]. The list was expanded using a snowball system, which led to 83 experts being contacted and 35 interviews scheduled. Three interviews could not be further processed because of recording issues, so a final sample of 32 interviews remained for analysis. Expert demographics are given in Table I.

Interview Procedure: Because the research was conducted during the COVID-19 pandemic, face-to-face interviews were neither possible nor ethical [59], [60]. Two researchers, aware of possible problems of lexicality [61], conducted semistructured interviews in English ($n = 8$) and German ($n = 24$) as unobtrusively and nondirectedly as possible [62] through a private Jitsi Meet [63] instance between October 2020 and January 2021, recording them with OBS Studio [64]. The interviews included questions about information categories of data processing relevant to risk communication, potential consequences of data processing, and means of mitigating potentially negative consequences (see Appendix B of the supplementary material). Each recording was assigned a random interview ID, indicating from which interview the excerpts originated. Translated excerpts are indicated by a T in the results section. The ethics board of the relevant institution approved the interview guide and anonymization procedure.

Interview Analysis: For the analysis of the interviews, we applied Mayring's SQC analysis method [56]. The first part of the interviews (about relevant aspects of the processing of personal data) was used to validate and expand the taxonomy resulting from the previous analysis. The interviews were coded in three steps.

1. Two researchers conducted a simultaneous deductive and inductive analysis of 37% of the material.
2. After a discussion, the coders expanded the codebook with inductive codes and recoded the material based on the expanded codebook, resulting in a Cohen's Kappa of 0.75.
3. Subsequently, the coders analyzed and discussed the remaining documents.

The interviews were used to investigate potential negative or unexpected consequences and their relationships to the established information categories of processing personal data. Researchers coded experts' opinions on whether adverse effects had to be understood concerning specific scenarios or across different contexts deductively as *general*, *specific,* or *both*.

Deductive codes for potentially adverse consequences were based on the systematization of negative consequences by Drackert [65] resulting from data processing and consequences named by the Article 29 Working Party [66]. Consequences not covered by the codebook were coded inductively, requiring specific examples from research or reporting and definitions based on relevant literature. Specific contexts explicitly related to consequences identified by experts were also coded inductively. Researchers determined a particular real-life example for each consequence in each identified context. Consequences mentioned by fewer than two experts were discarded. Segments about consequences were examined regarding explicit references to their causes (i.e., other consequences or risk-relevant information categories already included).

For coding, the aforementioned three steps were reapplied to the second part of the interviews.

Those referring to negative consequences in the processing of personal data resulted in a Cohen's Kappa of 0.87 after 37% of the material had been coded. We visualized the resulting connections between consequences and risk-relevant information categories to infer structural information.

One researcher coded the segments about risk mitigation inductively to identify specific strategies and responsible actors. Afterward, the authors discussed the results and assigned each strategy to an inductive category.

Results

This section details the results of our studies. First, we briefly summarize the information categories resulting from the GDPR SQC analysis. Second, relevant information categories resulting from the expert interviews are described. Then, we detail the results of applying the *Contextual Model for Perceived Privacy Risk* during analysis, presenting relations between events and consequences in processing personal data. The last section highlights potential guarding strategies against negative consequences that data subjects, data controllers, and governments can implement.

Identifying Relevant Information Categories for Risk Communication in the Processing of Personal Data The SQC analysis yielded five main categories: 1. data subject, 2. data type, 3. data processing, 4. processing purpose, and 5. security, with a total of 45 subcategories (see Appendix A, Table VIII, in the supplementary material).

Validating and Expanding the Taxonomy All categories identified in the SQC, except data alignment, had at least one expert mention. In total, 22 inductive subcategories (italicized in Appendix A, Table VIII, in the supplementary material) were added to the taxonomy. However, some codes, such as information about controllers ($n = 23$), applicable legal framework ($n = 22$), data subjects' rights ($n = 18$), and the responsible data protection authority ($n = 5$), were kept as codes for further analysis but not included in the taxonomy, as they were implicitly represented by the existing codes or followed logically from the object of study. These articles set forth which information has to be provided to the data subjects regarding the processing of their personal data in general. The following sections detail the results according to the defined categories.

Data Subject: The legal term data subject was associated with issues regarding comprehensibility: "I find it a bit … well, as a lawyer, I can imagine what it might mean, but as a normal user, it is a bit … scary, possibly even confusing" (85T).

Clarifying the meaning of the category as an umbrella term for risk-relevant information categories related to data subjects led to experts naming people with mental or physical impairments ($n = 6$) and children ($n = 16$) as vulnerable groups within data subjects. The legal status of children was frequently emphasized, "as a group treated separately in the GDPR" (07T) mainly because

> they do not have the age to make legally valid decisions, so to say. So, they're still under the responsibility of some other adults. And of course, that can be very vulnerable …. (51)

This special protection was considered especially relevant because "advertising to children, in general, can have very adverse effects" (08).

The category of vulnerable data subjects was considered crucial, and it was mentioned by 19 experts. The concept of vulnerability, however, does not necessarily apply only to specific groups of people as one expert explained.

> People often talk about the existence of vulnerable data subjects. In my opinion, this is not correct because every data subject is vulnerable. However, everyone is vulnerable in different situations. One is an expert in one area and another in another. But everyone has a weak spot somewhere, and many procedures aim to find and exploit these weak spots in a targeted manner. I think that is a great danger. … At the end of the day, I would say that it is often irrelevant, at least in terms of the technology used, whether it is advertising for elections or products. It is always the case that attempts are made to find certain trigger points of the people concerned, who are usually unaware of it …. (32T)

Especially nonlegal experts ($n = 11$) also mentioned collective behavior as relevant for individual privacy. Users could be targeted through inference of personal data based on data collected from other data subjects.

> … through that, they can, for example, find out that I might have a propensity for a gambling addiction, … even though I did not know it myself, and that they can target gambling ads,

> right? That is an example where my best interest and the best interest of the gambling company are not in alignment. I think that is one of the causes of these harms. (08)

Notably, the "nature of the data or the definition of the context" (14T) influences the data subjects' societal roles and their decision making. It also affects data-processing purposes, with experts expressing difficulty separating the context from the data-processing purposes. Apart from these collective and contextual aspects, however, interview participants stated that they were not sure about communicating information about the category *data subject* due to potential comprehensibility issues.

Data Type: Regarding data types, most experts recommended the categories of sensitive data ($n = 29$), health data ($n = 17$), and biometric data ($n = 16$) for risk communication. Additional information categories, such as financial information, were considered personal but not sensitive data due to the GDPR's clear definition of the latter (cf., Art. 9, Rec. 10 GDPR).

Both legal and nonlegal experts ($n = 12$) considered location data necessary for communication.

> If I have all the location data, I can create a movement profile for a person. Then the location data in its entirety becomes a kind of sensitive data. (33T)

Furthermore, location data serve as the basis for creating "encounter data, i.e., to what extent do people encounter each other" (81T).

Participants criticized the GDPR as it

> has these classes of special data or sensitive data on how you would look at it, which is complete nonsense. So, if I was a homosexual male, that's a protected characteristic, but my geolocation data that puts me in gay bars are not. I mean, it's just nonsense. (45)

Four experts explicitly drew the concept of a fixed set of sensitive data into question. While the distinction can be helpful, they saw data as not inherently risky as the risk depends on the actors and the processing purpose. For instance, health data are not risky when processed by a doctor. However, in the hands of other data controllers, the data might put data subjects at risk. Thus,

> you would probably have to differentiate between this internal social relationship, between the person concerned and the controller …, from the same data in relation to third parties. … And that is why [sensitive data types] are rather system-relative. That is, relative to the social and technical constellation. (04T)

Moreover, individual data points are frequently aggregated and merged with other data to gain new insights and increase informational and monetary value. Therefore, any approach that focuses heavily on data categories is viewed critically and would only make sense for processing raw, nonaggregated data, which is not the case when inference methods are applied during processing.

Experts without legal background ($n = 7$) considered unique identifiers relevant because this data type facilitates the deterministic combination of datasets, granting controllers access to previously unknown information about data subjects without employing probabilistic methods, such as machine learning. However, experts also mentioned that apart from (alphanumeric) identifiers, readily available personal data (e.g., name, surname, telephone number, address, or email), less available personal data (e.g., bank account number, birth date, or ID card number), or biometric data (e.g., voice recordings or user behavior) could also be used for data combination.

Data Processing: Data collection and storage are fundamental for all data-processing operations. Due to a lack of alternatives, data subjects "are excluded from important social and economic processes if they want to avoid or refuse data collection" (15T). Data collection was seen as a targeted acquisition of data ($n = 6$). However, two experts differentiated "between provided, observed, and inferred data" (54). They emphasized that data subjects should be informed if the data collection is unapparent, especially when recording or monitoring. For instance,

> in the smart home, the activation of a light switch already generates data—that is not so clear to many. In highly automated driving, every steering movement generates data. In other words: very banal data processing takes place here, which no longer appears on the device. (05T)

Furthermore, cookies and other more invasive means of tracking need to be considered, such as ultrasound signals or biometric technologies, often collected through smartphones. Therefore,

> not only do we disclose information about ourselves, such as account information or dates

> of birth or other such things, but we also intentionally or unintentionally disclose data about ourselves by tracking the place where we are. (98T)

Although collection makes data initially available, processing requires access to stored data. Therefore, the storage period ($n = 15$) and storage location ($n = 14$) were also considered relevant, which may make alteration ($n = 5$) or data erasure ($n = 13$) more difficult. The latter two constitute essential data protection rights guaranteed by Art. 16 and 17 of the GDPR [1], which allow data subjects to request the rectification of inaccurate personal data or—under certain circumstances—the deletion of personal data.

Apart from the categories data collection ($n = 17$) and storage ($n = 22$), experts mainly mentioned two processing operations which they would inform data subjects about: inference ($n = 18$) and disclosure ($n = 26$). Various mechanisms are meant with inference, described as "individual pieces of information [that] are puzzled together by machines, and then new information is generated" (20T), which is especially problematic "when data sets are brought together and then used for forecasting or classification purposes." Hence, "at the time of data collection, it is not at all clear what information will be generated in the end" (03T).

Apart from data collection and statistical techniques, however, inference can also be achieved through behavioral feedback.

> So as an advertiser, I know that if this group or this cluster responds to my advertising, that I have done the right thing and can perhaps act even more specifically … and get more information. It is a kind of loop. It is a cycle that leads to increasingly sophisticated personality profiles of individuals or groups. (98T)

Thus, collected and inferred information is "put together into very complex personality profiles," based on which "negative decisions can, of course, then be made for me, such as that I pay a higher price for future purchases on a platform" (91T).

These processes also relate to automated decision making (ADM, $n = 8$). Here, a decision is deduced (or inferred) based on existing data. ADM is similar to inference as data subjects often also do not know if they are subject to automated decisions and on what data such decisions are based. However, it is potentially more severe as these decisions (or inferences) generally entail more significant effects.

However, the most mentioned data-processing operation was data disclosure. Most experts ($n = 26$) assign particular importance to the questions whether and with whom data are shared. Thus, for a risk evaluation, both the controllers sharing and the controllers or processors receiving the data are relevant, and according to the experts, "every transfer of data to third parties should be marked" (91T).

Many experts also mentioned anonymization ($n = 11$) and pseudonymization ($n = 9$) operations. In the case of full anonymization

> once this depersonalization action takes place …, all of the GDPR and public law instruments granted rights, they become quite futile. (96)

Some experts perceived these operations as preserving privacy ($n = 5$). In contrast, 13 experts highlighted that these techniques are not a definitive solution to the issue of personal data processing because both anonymization and pseudonymization can be reversed based on inference techniques, mainly when multiple datasets are aggregated or linked.

According to our interview results, processing should always be disclosed, especially if it is unexpected by the data subject. That emphasizes the importance of inference as it can be used to access previously unknown and thus unexpected data. Disclosure makes it hard to track what happens with the data afterwards and increases the chance of unexpected processing operations with unforeseen purposes.

Processing Purposes: Many experts ($n = 14$) saw a change in data-processing purposes as most relevant for data subjects—a practice that, according to the GDPR, is illegal if done without notice and if it does not comply with further data protection requirements [Art. 6 (4), Art. 13 (3), Art. 14 (4)].

In terms of phrasing, overly specific purposes were viewed as impractical. Meanwhile, general purposes were criticized as "the purposes with the largest potential of abuse" (81T). *Service optimization* is an example of a general purpose, potentially encompassing improvements to the customer experience and increasing the company's profit margin. Specific types of data collected can be paired with the associated purposes of collection to provide accountability for the controller as

> you want to know whether I click on a product or not, but why is it important for you to know my gender or whatever? And then you have to do some justification work. (05T)

According to some experts ($n = 6$), data subjects should also be informed about commercial and business purposes (i.e., marketing or the sale of data) to indicate to users whether free usage of services directly relates to them "actually being commodified right now" (35). Financial motifs are relevant because they increase the likelihood that personal data are used to generate profits. To this end, additional information that data subjects might be unwilling to disclose might be collected or inferred, or data might be sold. That information increases the chances of data being used for not-agreed-upon purposes. Three experts additionally suggested disclosing the services' business model as "an important signal for … a user," which would

> support the plausibility of the data protection information, because if a company cannot state how it finances itself or how it finances its service, but at the same time claims that data protection is infinitely important to them, then of course that is not so plausible. (91T)

Security: Due to low inter-rater reliability, security served as an explorative category, and after more in-depth analysis, was eventually considered a modality of data processing as "you can process data in a way that is secure or very insecure" (14T). Thus, security was included in the final taxonomy but will not be discussed as elaborately as the other main information categories.

Security mechanisms guard against unintended data processing but not against other sources of risk, such as purpose change or data disclosure. Law or computer science experts were explicitly asked whether and how they would prioritize different security aspects. Although some chose technical measures over organizational ones ($n = 4$) or the other way around ($n = 3$), the majority ($n = 11$) of experts argued for the interdependence of both factors, refusing to prioritize either one.

All experts noted technical measures that could be subdivided into software- and hardware-based interventions. They ($n = 9$) referred to encryption when mentioning software-based security mechanisms (specifically Transport Layer Security and end-to-end encryption). As a relevant information category concerning hardware, they mainly considered the location of data storage ($n = 14$). Some saw this as relevant because "if [data] are stored in third countries," they might not be under "an adequate level of data protection" (21T). In contrast, others pointed out that "it does not do me any good if I have the data in a storage location that is supposedly safe but lying openly on the internet within Europe" (61T).

In total, 17 experts also mentioned that access control could limit the number of individuals and the purposes involved in processing, including logging data access to reconstruct it in case of a leak.

Organizational measures were named less often ($n = 20$) but considered just as relevant by security experts because "the weakest link in the chain is the human" (14T). They included education and training against social engineering and developing safe data handling practices ($n = 11$). Five experts also mentioned certification as relevant information for data subjects to ensure that at least minimal security standards are applied to the processing of their data.

Hazards in the Processing of Personal Data

Events Leading to Consequences in the Processing of Personal Data: At the top, Fig. 3 indicates that data collection and storage are crucial events for risk formation, on which all further processing operations are based—as one follows the arrows from the top left to the top right. This first row of events in the processing of personal data shows that specific data-processing operations are the main causes of negative or unexpected consequences because all further consequences follow below. For example, new information can be derived (leading to the consequence "Emergence of Information") using stored data based on already existing data (inference) or by combining different data sets (data combination).

Another relevant hazard is data disclosure, which might result in a transgression of contexts and a change in processing purposes. Processing other purposes than those listed is also considered as a hazard. Data type and subject categories are absent from these risk relations. Instead, Fig. 3 shows that other "upstream" consequences can also cause adverse effects. The next section further specifies the resulting relationships.

Consequences of Processing Personal Data: When asked about the context dependence of adverse consequences, a third of experts said they were either general ($n = 6$) or specific ($n = 5$). The

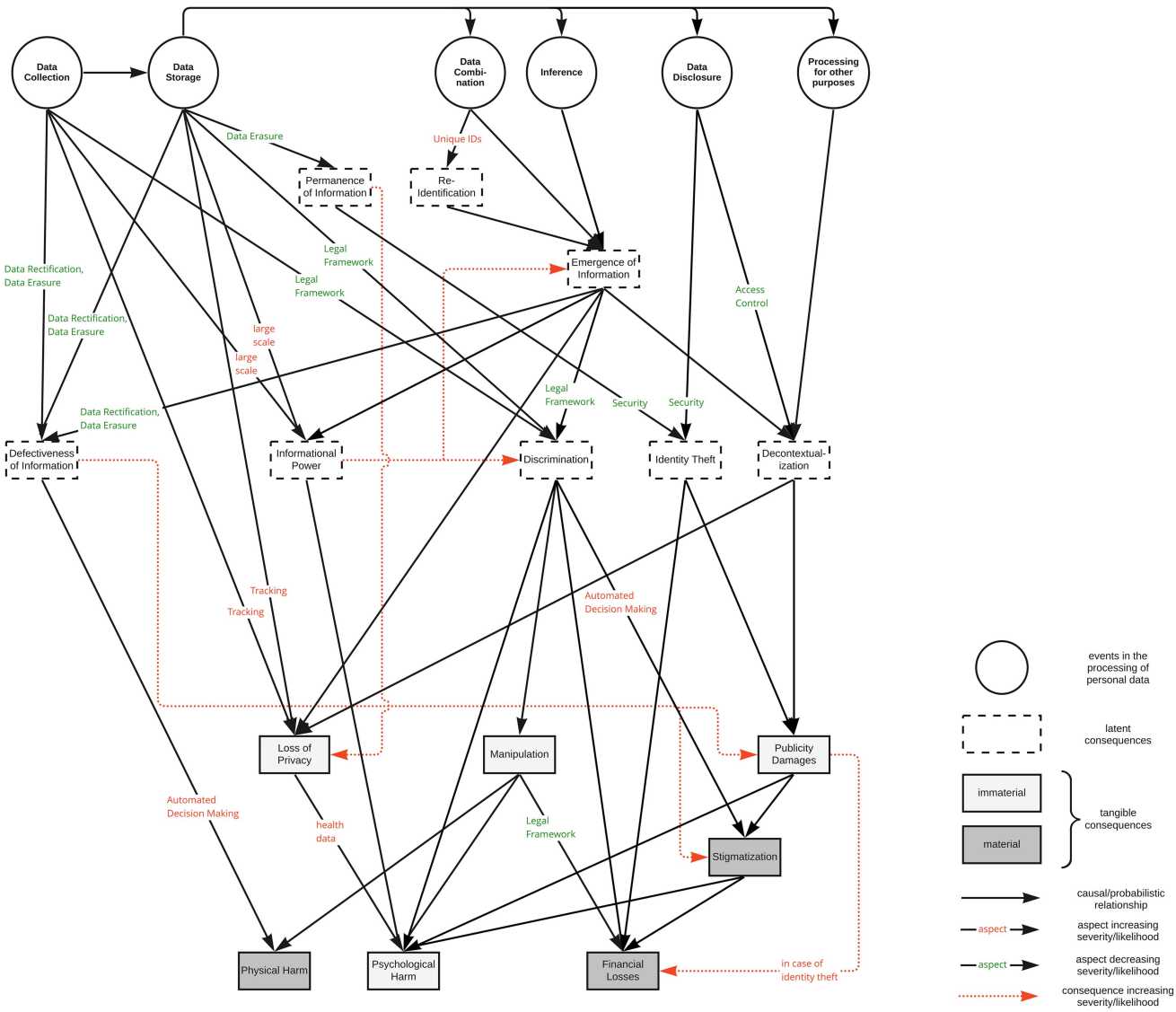


Fig. 3. Model based on experts' answers when asked about negative or unexpected consequences resulting from data processing, their origin, and the information categories influencing their likelihood or severity. At the top, multiple data-processing operations that experts linked to consequences are shown in circles. Below, latent consequences are connected to material and immaterial tangible consequences.

remaining two-thirds ($n = 20$) stated that consequences could apply across and within contexts (with financial, health, and work contexts being mentioned repeatedly). Plotting and ordering of the relations between events and the identified negative consequences revealed a structure of latent and tangible consequences shown in Fig. 3. For detailed information about causes, potential modulations, explanations, and examples for latent and tangible adverse consequences, see the supplementary material, Appendix C, Tables IX and X, respectively.

The latent consequences listed in Table II are first-order consequences that are not immediately known to or experienced by the data subject. Contexts are rarely mentioned concerning latent consequences, indicating that they are not bound to specific actors or data types. Meanwhile, *tangible consequences* can be directly experienced by data subjects, manifesting as a result of latent consequences, as given in Table III. They can be subdivided into material and immaterial consequences. Experts also referred to tangible consequences in specific contexts more frequently.

Guarding Strategies against Negative or Unexpected Consequences Table IV shows that experts identified three actors responsible for preventing or reducing the severity of

TABLE II
LATENT CONSEQUENCES OF DATA PROCESSING BASED ON EXPERT INTERVIEWS

Consequences	General Definition	Causes	Potential Modulators	*N*	Contexts
Informational Power	Concentration of power as a result of information imbalances that can be used as an instrument for state and private power [64].	Data Collection & Data Storage, Emergence of Information	Large Scale Processing	17	-
Decontextualization	Cross-sectoral usage of personal information characterized by surprising effects and lack of controllability [64].	Emergence of Information, Data Disclosure, Processing for other Purposes	Access Control	14	-
Emergence of Information	Possibility of automatically drawing new, previously unavailable conclusions from various pieces of information [64].	Data Combination, Reidentification, Inference	Informational Power	12	health (*n* = 6)
Discrimination	Categorical treatment of a person and an associated negative evaluation [64].	Data Collection & Storage, Emergence of Information	Legal Framework, Informational Power	9	-
Reidentification*	Tracing anonymized data back to the data subject to whom it relates [66].	Data Combination	Unique Identifiers	6	-
Identity Theft	Acquisition of personal data or documents by an attacker who uses this information to pass themselves off as the victim.	Permanence of Information, Data Disclosure	Security	6	-
Defectiveness of Information	Inaccuracy of personal information [64].	Data Collection & Storage, Emergence of Information	Data Rectification, Data Erasure	4	-
Permanence of Information	Adverse effects that result from the long-term availability of information, influenced by factors such as digitization, storage capacity, decentralization, and redundancy of storage on the internet [64].	Data Storage	Data Erasure	3	-

TABLE III
TANGIBLE CONSEQUENCES OF DATA PROCESSING BASED ON EXPERT INTERVIEWS

Type	Consequence	General Definition	Causes	Potential Modulators	*N*	Contexts
Material	Financial Losses	Loss of money or access to financial resources.	Stigmatization, Discrimination, Identity Theft, Manipulation	Publicity Damages, Legal Framework	23	financial (n = 23)
	Stigmatization	Attribution of characteristics to a person that forms the basis for avoiding or excluding them [64].	Discrimination, Publicity Damages	Defectiveness of Information, ADM	17	work (n = 9)
	Physical harm	Injury or damage to the physical health of people [67].	Manipulation, Defectiveness of Information	ADM	5	health (n = 2)
Immaterial	Loss of Privacy*	Perceived violation of existing (contextual) informational norms, sometimes connected with a perceived loss of control.	Data Collection & Storage, Data Disclosure, Emergence of Information	Tracking, Permanence of Information, Access Control	22	Work (n = 2)
	Psychological harm*	Injury or damage to the psychological health of people [67].	Loss of Privacy, Informational Power, Manipulation, Publicity Damages, Stigmatization	Health (genomic) data	11	health (*n* = 3)
	Manipulation*	Imposing a hidden or covert influence on another person's decision-making [68].	Discrimination		8	financial (*n* = 5)
	Publicity damages	Harms occurring without additional circumstances outside the person concerned, i.e., already with the publicity of the information [64].	Decontextualization, Identity Theft	Defectiveness of Information	7	work (*n* = 7)

TABLE IV
SPECIFIC GUARDING STRATEGIES AGAINST NEGATIVE OR UNEXPECTED CONSEQUENCES DIFFERENTIATED BY RESPONSIBLE ACTORS AND STRATEGY

Responsible Actors	*N*	Strategy Category	*n*	Specific Strategies
Data Controllers	21	Security Measures	17	1. Technical: encryption (*n* = 7), ensuring that hardware is safe and protected against unauthorized access (*n* = 5), access control/identity management (*n* = 4)
				2. Organizational: education and Training, e.g., for resisting social engineering (*n* = 7)
				3. Certification for technical (*n*=3) and organizational measures (*n* = 2)
		Processing Information for Nonexperts	8	1. Possible consequences and countermeasures (*n* = 3)
				2. Clear communication of purposes (*n* = 2)
		Processing Guidelines	6	1. Purpose Limitation (*n* = 6)
				2. Data Minimization (*n* = 2)
				3. Making deletion of data more accessible (*n* = 2)
		Design Guidelines	3	Involve data subjects in technology design process either through communication or User-Centered Design (*n* = 3).
Governments	8	Enforcement	4	Enforcement of existing regulations (*n* = 4).
		Promote Awareness	3	Promote data privacy awareness in the society (*n* = 3).
Data Subjects	7	Incorporate Preventive Measures	4	1. Develop personal privacy practices (*n* = 3).
				2. (Non-)Use of specific services to reduce the amount of data generated (*n* = 2).
		Exercise Rights	3	Exercise use of user-rights, such as right to deletion, rectification, or restriction of processing (*n* = 3).

consequences: controllers, governments, and data subjects. Noticeably, they mainly provided rather abstract answers, indicating the difficulty of reducing the likelihood or severity of consequences.

The distribution of answers indicates that data controllers can implement most strategies as they design the data collection and sharing systems. That includes security measures, which prevent unauthorized access to data and thus mostly guard against identity theft and unwanted decontextualization. However, controllers can also reduce risks by limiting the data collected and the associated purposes. Moreover, controllers should better inform data subjects by providing information about links between purposes and collected data types, potential consequences resulting from processing personal data, and measures taken to mitigate these risks. These steps would require data controllers not just to rework their privacy notices to provide this information in an easily understandable format but also to offer simple ways for users to exercise their data protection rights, especially when it comes to the deletion of data.

As for governments, experts mainly saw them as responsible for guaranteeing individual data protection rights by enforcing the existing regulation and fostering privacy awareness.

Although data subjects have the fewest options to reduce risks, they can prevent data collection by avoiding specific services or using privacy-preserving technologies. Furthermore, data subjects can mitigate risks by exercising their data protection rights, including the rights to rectification, erasure, and restriction of processing as guaranteed by the GDPR [1, Art. 16, Art. 17, and Art. 18], assuming that the controllers act compliant to the regulation.

LIMITATIONS

Before discussing the results, it is essential to highlight this study's methodological and conceptual limitations. Conducting expert interviews online comes with limitations for data collection because of poor internet connections or varying capabilities in using information technologies [59]. The generalizability of the study is limited because it was based on German GDPR commentaries, and the expert sample was biased toward European and German-speaking academics with legal backgrounds. As a result of these factors, the findings might be more GDPR-congruent.

Conceptually, the most fundamental limitation concerning our approach is the notice and choice framework. Risk communication within privacy notices can help users make informed consent

decisions. At the same time, it is essential to acknowledge that it cannot fix structural issues, such as a lack of choice regarding certain service providers. Another issue is the lack of comparability regarding different degrees of information disclosure due to the absence of standardizations. It cannot also fix conceptual problems inherent to the notice and choice framework, such as the assumptions that humans behave as rational actors in a consent decision and that these decisions only affect the individual data subject [70].

Another conceptual limitation of this article is the attempt to define a closed set of data types. This approach has repeatedly been questioned, as new information can constantly be gathered or inferred from existing data without conforming to previously established information categories [44]. Moreover, purposes are similarly difficult to categorize, as they can evolve and might thus escape predetermined taxonomies or standardization efforts without constant actualizations.

The event-based approach of the perceived risk model and the contextual model for perceived privacy risk are also restricted as they include only consequences that result from specific data-processing events. Although this approach is appropriate when examining the consequences of data processing (which assumes that events like collection and storage occur), it is essential to acknowledge that data subjects can also experience negative consequences resulting from the absence of data processing [71], [72].

Discussion

Relevant Information Categories for Risk Communication in the Processing of Personal Data Our results confirm and expand our initial taxonomy of relevant information categories for risk communication regarding the processing of personal data. Except for data alignment, all deductive categories in the codebook were mentioned by at least one expert, which allows for two possible conclusions: either the GDPR [1] and the Art. 29 Working Party publications [34], [73], [66] achieve sufficient coverage of risk-relevant information categories in the processing of personal data as identified by experts. Alternatively, these texts significantly influenced the interviewed experts.

Interestingly, the GDPR does not directly address the issue of inference. Instead, it mentions only related data-processing practices, such as profiling, scoring, data combination, or ADM. That fact indicates a significant blind spot concerning risk, even questioning the understanding of sensitive data. Inference techniques offer the possibility of deducing sensitive data based on nonsensitive data [44].

Another set of relevant information categories, repeatedly mentioned by experts and not explicitly mentioned by the GDPR, is commercial purposes, especially the sale of data. The sale of data can serve as an indication that data controllers commodify data subjects (or their personal data) and might entail more excessive data collection and sharing practices [74] and a violation of the data minimization principle from Art. 5 (1)(c) GDPR [1]. This principle posits that personal data shall be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed" (Art. 5 (1)(c) GDPR) [1]. Excessive and broad purposes pose additional risks. They could indicate that personal data might be processed for unexpected purposes, which would, in turn, violate the principle of purpose limitation [Art. 5 (1)(b)]. This principle states that personal data shall be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes."

Relations Between Relevant Information Categories and Potential Consequences Fig. 3 clarifies that data-processing operations play an integral role in forming consequences within the data controller and data subject relationship because they are equivalent to events within models of perceived risks. Data must be collected and stored (i.e., processed) in the first place, with more potential consequences arising from further processing operations, such as combination, inference, and data disclosure. We derive from this centrality of data-processing operations in risk creation that they should also play a central part in risk communication. Notably, our resulting set of risky processing operations shows similarities to Nissenbaum's distinction of sources of privacy risk, even though we arrived at it independently and in a partially inductive process. Those risky operations are data collection, storage, and monitoring (corresponding to monitoring and surveillance), data disclosure (dissemination and communication), and data combination, profiling, ADM, and scoring (aggregation and analysis) [42, pp. 191–230], [44].

Still, effective risk communication should not solely include risky processing operations but also the

associated potential consequences of processing personal data [75]. Here, the results suggest differentiating between latent and tangible consequences. Both levels are essential to consider because latent consequences mediate between data-processing practices and more specific tangible consequences. As latent consequences do not have to exist within specific contexts, they are much more abstract and thus more challenging to communicate.

On the other hand, tangible consequences are often part of an explicit warning system, in which displayed risks are usually specific and comprise a physical safety component or the possibility of financial loss [40], [76]. We argue that privacy risk communication should strive to include both kinds of consequences to enable data subjects to better understand which latent consequences cause more tangible consequences. Thus, the information aids data subjects in constructing a procedurally accurate and more complete mental model of personal privacy risk.

If applied this way, the proposed model can reconcile general and context-specific understandings of risks, conceptualizing them as different pathways leading to the formation of consequences. Since all data are usually collected and stored within a specific context, if the data are adequately secured and processed only for specific, agreed-upon purposes, the potential (tangible) consequences should not occur or arise only in the original context. However, as the results show, multiple data-processing practices allow for the transgression of contexts, either directly or through latent consequences related to decontextualization. They thereby challenge the contextual separation fundamental for Nissenbaum's understanding of privacy [42].

Today's information technology now permeates nearly all such contexts and allows for context-independent information storage and extensive data flows between actors, transgressing contexts and information categories within milliseconds [77]. Contextual information flows can thus be understood as nested within the overarching context of Big Data, which is "incompletely specified" [42, p. 135] and within which some latent consequences, such as decontextualization, can occur. If information is decontextualized, it becomes increasingly complex to anticipate what kind of other consequences might follow. If we assume that the data subject faces negative or unexpected (tangible) consequences, then these consequences always occur within a specified context, which can differ from the initial context.

Means of Avoiding or Mitigating Negative or Unexpected Consequences According to some of the interviewed experts, it seems more fitting to refer to data subjects metaphorically as data objects because their knowledge of potential risks resulting from processing their personal data is often as limited as their options for reducing them. Still, existing frameworks for risk communication state that warning notices should include meaningful options and instructions on how to avoid harm [78]. Thus, data subjects should be informed about (or even encouraged to exercise) their rights, especially as to erasure and rectification, and restriction of processing. However, Resnick [79] emphasized that data subjects are unlikely to take advantage of their rights unless the process is easy to access and complete.

According to experts, controllers are responsible for implementing privacy controls and preservations. In addition, controllers should inform data subjects about their processing practices in a way that provides them with an understanding of the processing and informs them about potential consequences arising from data processing and means of mitigating them. In addition, governments should educate citizens about their rights and enforce existing regulations.

Conclusion

What to Emphasize for Risk Communication in Privacy Notices? Integrating the detailed results discussed above, we can conclude what information should be emphasized in risk communication to data subjects. Table V summarizes relevant categories for risk communication. They are ranked according to relevance based on interview document frequencies and theoretical considerations. These categories inform data subjects about potential sources of risks and meet many of the information criteria defined by the GDPR [80].

Considering the centrality of data processing for risk formation in the *Contextual Model for Perceived Privacy*, we suggest presenting data processing as the starting point for risk communication. Furthermore, risk communication should include relevant latent and tangible consequences to provide data subjects with adequate mental models of risk formation. Finally, to enable data subjects to

TABLE V
PROPOSED INFORMATION CATEGORIES FOR RISK COMMUNICATION, SORTED BY RELEVANCE WITHIN THEIR RESPECTIVE SUPER CATEGORIES

Super Category	Proposed Information Category
Latent Consequences	Informational Power
	Decontextualization
	Emergence of Information
	Discrimination
	Identity Theft
	Defectiveness of Information
	Permanence of Information
Tangible Consequences	Financial Losses
	Stigmatization
	Physical Harm
	Publicity Damages
	Manipulation
	Psychological Harm
Risky Data-Processing Operations	Profiling*
	Automated Decision Making*
	Scoring
	Behavioral Monitoring
	Systematic Monitoring of Data Subjects on a Large Scale
	Data Collection from Third Parties*/Data Combination
	Data Disclosure*
	Large Scale Processing of Personal Data
	Storage Period*
	Off-Device Storage
Risky Data Types	Sensitive Data*
	Genetic Data*
	Health Data*
	Biometric Data*
	Location Data
	Unique Identifiers
	Financial Information
Risky Processing Purposes	Data Sale*
	Excessive or Broad Purposes*
Data Subject Rights	Right to erasure*
	Right to rectification*
	Right to object*
	Choices*

Note: Categories that meet the information requirements in GDPR Art.13 and 14 are highlighted with an asterisk.

take preventive measures, we suggest more explicit communication of data protection rights, such as the right to revoke consent as already required by the GDPR, and more generally speaking, we suggest supporting data subjects by facilitating access to and easing understanding of privacy choices [81].

Contribution and Potential Applications Our article makes three significant theoretical contributions. First, the contextual model for perceived privacy risk expands existing models with context-dependent information, explaining events and consequences within and across contexts.

Second, we identify various negative consequences of personal data processing and link them to relevant events. As a result, we specify the model of perceived privacy risk for processing personal data, providing detailed information on hazards and consequences and potential actors and actions that could avoid or mitigate risk. This systematization can be the basis for a more fine-grained risk assessment based on usage practices. Data controllers can use it to design privacy-preserving technologies by identifying risky processing operations that are especially critical for user privacy and warrant the implementation of guarding strategies. Alternatively, data protection authorities may use the model to estimate potential harms that might result from data processing by various services, which might help them better protect individuals' data protection rights.

Third, we propose a set of information categories for privacy risk communication. Communication professionals can additionally use these for redesigning, highlighting, or visualizing relevant aspects of their privacy notices to transparently communicate potential consequences associated with specific processing operations and the measures taken to protect the users of a given service. This way, the specified model can inform users about risks and potential courses of action. It can thus serve as a valuable basis for privacy risk education. Generally, risk communication could effectively counter cognitive limitations in consent situations since it could address issues involved in individual decision-making behavior incongruent with attitudes about privacy [29].

Accordingly, communication professionals can support data subjects to create more accurate mental models, and improve risk perception and appropriate privacy-protective behavior. This will likely be reflected in product selection when services can be easily compared [22]. Data protection authorities could support these efforts by developing a standardized set of privacy icons for risk communication based on the proposed information categories. Widely communicating these risk-based information categories would benefit the common understanding of privacy risks and favor competition between providers, considering the European Commission's goal to become a "leading role model for a society empowered by data to make better decisions—in business and the public sector" [82, p. 1].

Further Research Since the formation of privacy risk perceptions and their role in decision making is not currently well understood [55], our model of risk formation can provide insight into differences and test user knowledge and conceptions about latent consequences by comparing it to models drafted by laypeople. Quantitative expert and user research accounting for varying perceptions of severity and likelihood of the various latent and tangible consequences is needed to understand further which consequences to prioritize in communication. We also call for more research into potentially risky purposes to identify those purposes that are defined in a broad enough sense to indicate potential abuse.

Considering that additions to the GDPR and large-scale self-imposed risk communication by controllers are unlikely to happen soon, we advocate combining our results with large language models for scalable analysis of privacy notices to automate the identification and presentation of risky information categories. That goal could be implemented as a browser plugin, similar to the one suggested by Harkous et al. [83], which would provide users with an indication of several risk-relevant aspects of data processing. Such an approach could directly be used to analyze risk communication's optimal presentation and effectiveness.

Acknowledgment

The authors would like to thank Prof. Christiane Wendehorst, Prof. Martina Angela Sasse, Prof. Max von Grafenstein, Dr. Lena Ulbricht, Dr. Michelle Christensen, Dr. Jörg Pohle, Dr. Stefan Ullrich, Dr. Arianna Rossi, Dr. Johanna Schäwel, Dr. Frank Pallas, Dr. Alain Couillault, Linda Bienemann, Sebastian Eschrich, Karolina Iwanska, Sonja Waldgruber, Florian Glatzner, Frank Ingenrieth,

Anna Schenk, Frederick Richter, Ailo Krogh Ravna, Daniel Guagnin, Richard Huber, Joss Langford, and Susanne Seibold, and all other participants who preferred to stay anonymous, for participating in our interviews. Additionally, we want to thank Nieke Wagner, Malte Mackensen, and Niklas von Kalckreuth for their assistance in data collection for the content analysis. They would also like to thank Tetiana Shportak for creating a contact list for our expert interviews as well as Julian Boltz and Lisa Völzmann for supporting the interview transcription efforts.

This work was supported in part by the Federal Ministry of Education and Research (BMBF) under Grant 16DII131, and in part by the Open Access Publication Fund of the Weizenbaum Institute for the Networked Society, Berlin.

References

[1] European Parliament and Council of Europe, "Regulation (EU) 2016/679 of the European Parliament and of the Council of Europe 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 (GDPR)," May 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj
[2] P. Voigt and A. Von Dem Bussche, *The EU General Data Protection Regulation*. Berlin, Germany: Springer, 2018.
[3] C. Utz, M. Degeling, S. Fahl, F. Schaub, and T. Holz, "(Un)informed consent: Studying GDPR consent notices in the field," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2019, pp. 973–990.
[4] D. Solove, "Introduction: Privacy self-management and the consent dilemma," *Harvard Law Rev.*, vol. 126, no. 7, pp. 1880–1903, May 2013. [Online]. Available: https://ssrn.com/abstract=2171018
[5] C. Matte, N. Bielova, and C. Santos, "Do cookie banners respect my choice?: Measuring legal compliance of banners from IAB Europe's transparency and consent framework," Feb. 2020. [Online]. Available: https://arxiv.org/pdf/1911.09964.pdf
[6] Q. Weinzierl, "Dark Patterns als herausforderung für das Recht. Rechtlicher Schutz vor der ausnutzung von Verhaltensanomalien," *Neue Zeitschrift für Verwaltungsrecht – Extra.*, vol. 15, pp. 1–11, Aug. 2020. [Online]. Available: https://content.beck.de/NVwZ/NVwZ-Extra_2020_15.pdf
[7] B. Fabian, T. Ermakova, and T. Lentz, "Large-scale readability analysis of privacy policies," in *Proc. Int. Conf. Web Intell.*, 2017, pp. 18–25, doi: 10.1145/3106426.3106427.
[8] A. M. McDonald and L. F. Cranor, "The cost of reading privacy policies," *I/S: J. Law Policy Inf. Soc.*, vol. 4, no. 3, pp. 543–568, 2008. [Online]. Available: https://kb.osu.edu/server/api/core/bitstreams/a9510be5-b51e-526d-aea3-8e9636bc00cd/content
[9] J. A. Obar and A. Oeldorf-Hirsch, "The biggest lie on the Internet: Ignoring the privacy policies and terms of service policies of social networking services," *Inf., Commun. Soc.*, vol. 23, no. 1, pp. 128–147, 2016, doi: 10.1080/1369118X.2018.1486870.
[10] H. A. Simon, "Bounded rationality," in *The New Palgrave. A Dictionary of Economics*. London, UK: Palgrave Macmillan, 2008, pp. 1–4.
[11] D. Kahneman, "A perspective on judgement and choice: Mapping bounded rationality," *Amer. Psychol.*, vol. 58, no. 9, pp. 697–720, 2003, doi: 10.1037/0003-066X.58.9.697.
[12] A. E. Waldeman, "Cognitive biases, dark patterns, and the 'privacy paradox'," *Current Opin. Psychol.*, vol. 31, pp. 105–109, Feb. 2020, doi: 10.1016/j.copsyc.2019.08.025.
[13] R. Moll, S. Pieschl, and R. Bromme, "Trust into collective privacy? The role of subjective theories for self-disclosure in online communication," *Societies*, vol. 4, pp. 770–784, 2014, doi: 10.3390/soc4040770.
[14] A. Acquisti and J. Grossklags, "Privacy and rationality in individual decision making," *IEEE Secur. Privacy*, vol. 3, no. 1, pp. 26–33, Jan./Feb. 2005, doi: 10.1109/MSP.2005.22.
[15] P. A. Norberg, D. R. Horne, and D. A. Horne, "The privacy paradox: Personal information disclosure intentions versus behaviors," *J. Consum. Affairs*, vol. 41, no. 1, pp. 100–126, 2007, doi: 10.1111/j.1745-6606.2006.00070.x.
[16] S. Trepte, T. Dienlin, and L. Reinecke, "Risky behaviors: How online experiences influence privacy behaviors," in *Von Der Gutenberg-Galaxis Zur Google-Galaxis. Alte und Neue Grenzvermessungen Nach 50 Jahren DGPuK*, B. Stark, O. Quiring, and N. Jackob, Eds. Munich, Germany: UVK, 2014, pp. 225–244.
[17] S. Passera, "Beyond the wall of text: How information design can make contracts user-friendly," in *Design, User Experience, and Usability: User and Interactions*, A. Marcus, Ed., Berlin, Germany: Springer, 2015, pp. 341–352.
[18] M. Tabassum, T. Kosinski, and H. R. Lipford, "'I don't own the data': End user perceptions of smart home device data practice and risks," in *Proc. 15th Symp. Usable Privacy Secur.*, 2019, pp. 435–450.
[19] J. Lau, B. Zimmermann, and F. Schaub, "Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers," *Proc. ACM Human-Comput. Interaction*, vol. 2, no. 102, pp. 1–31, 2018, doi: 10.1145/3274371.
[20] A. Deuker, "Addressing the privacy paradox by expanded privacy awareness-the example of context-aware services," in *Privacy and Identity Management for Life*, J. Camenisch, S. Fischer-Hübner, and K. Rannenberg, Eds. Berlin, Germany: Springer, 2010, pp. 275–283.
[21] K.-P. L. Vu, V. Chambers, B. Creekmur, D. Cho, and R. W. Proctor, "Influence of the privacy bird user agent on user trust of different web sites," *Comput. Ind.*, vol. 61, no. 4, pp. 311–317, May 2010, doi: 10.1016/j.compind.2009.12.001.

[22] G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A 'nutrition label' for privacy," in *Proc. 5th Symp. Usable Privacy Security*, 2009, pp. 1–12, doi: 10.1145/1572532.1572538.
[23] S. Barth, D. Ionita, M. D. T. de Jong, P. H. Hartel, and M. Junger, "Privacy rating: A user-centered approach for visualizing data handling practices of online services," *IEEE Trans. Prof. Commun.*, vol. 64, no. 4, pp. 354–373, Dec. 2021, doi: 10.1109/TPC.2021.3110617.
[24] C. Graf, C. Hochleitner, P. Wolkerstorfer, J. Angulo, S. Fischer-Hübner, and E. Wästlund, "Final HCI research report," May 2011. [Online]. Available: http://primelife.ercim.eu/results/documents/148-415d
[25] A. Rossi and M. Palmirani, "DaPIS: An ontology-based data protection icon set," in *Knowledge of the Law in the Big Data Age*, G. Peruginelli and S. Faro, Eds. Amsterdam, The Netherlands: IOS Press, 2019, pp. 181–195, doi: 10.3233/FAIA190020.
[26] M. Meldau, "Iconset for data-privacy declarations v0.1." [Online]. Available: https://netzpolitik.org/wp-upload/data-privacy-icons-v01.pdf
[27] B. Moskowitz and A. Raskin, "Privacy icons," 2011. [Online]. Available: https://wiki.mozilla.org/Privacy_Icons
[28] F. Thouverain, M. Glatthaar, J. Hotz, C. Ettlinger, and M. Tschudin, "Privacy icons: Transparenz auf einen blick," Jusletter 30, Nov. 2020. [Online]. Available: https://privacy-icons.ch/wp-content/uploads/2020/12/jusletter_privacyicons.pdf
[29] Z. Efroni, J. Metzger, L. Mischau, and M. Schirmbeck, "Privacy icons: A risk-based approach to visualisation of data processing," *EDPL*, vol. 5, pp. 352–366, 2019, doi: 10.21552/edpl/2019/3/9.
[30] A. Rossi and M. Palmirani, "What's in an icon? Promises and pitfalls of data protection iconography," in *Data Protection and Privacy: Data Protection and Democracy*, D. Hallinan, R. Leenes, S. Gutwirth, and P. De Hert, Eds. Oxford, UK: Hart Publishing, 2020, pp. 59–92.
[31] K. R. Laughery and D. Paige-Smith, "Explicit information in warnings," in *Handbook of Warnings*. Mahwah, NJ, USA: Lawrence Erlbaum, 2006, pp. 419–428.
[32] D. K. Mulligan, C. Koopman, and N. Doty, "Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy," *Trans. Roy. Soc. A: Math., Phys. Eng. Sci.*, vol. 374, pp. 1–17, Dec. 2016.
[33] ISO, *ISO Guide 73:2009(en) Risk Management — Vocabulary*, 2009. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:guide:73:ed-1:v1:en
[34] Article 29 Data Protection Working Party, "Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is 'likely to result in a high risk' for the purposes of Regulation 2016/679," Apr. 2017. [Online]. Available: https://www.datenschutz-grundverordnung.eu/wp-content/uploads/2017/07/wp248_enpdf.pdf
[35] Y. Li, "Theories in online information privacy research: A critical review and an integrated framework," *Decis. Support Syst.*, vol. 54, no. 1, pp. 471–481, Dec. 2012, doi: 10.1016/j.dss.2012.06.010.
[36] S. Glover and I. Benbasat, "A comprehensive model of perceived risk of E-commerce transactions," *Int. J. Electron. Commerce*, vol. 15, no. 2, pp. 47–78, 2010. [Online]. Available: https://www.jstor.org/stable/27919912
[37] F. Kehr, T. Kowatsch, D. Wentzel, and E. Fleisch, "Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus," *Inf. Syst. J.*, vol. 25, no. 6, pp. 607–635, Mar. 2015, doi: 10.1111/isj.12062.
[38] T. Wang, T. D. Duong, and C. C. Chen, "Intention to disclose personal information via mobile applications: A privacy calculus perspective," *Int. J. Inf. Manage.*, vol. 36, no. 4, pp. 531–542, Aug. 2016, doi: 10.1016/j.ijinfomgt.2016.03.003.
[39] T. Jakobi, S. Patil, D. Randall, G. Stevens, and V. Wulf, "It is about what they could do with the data: A user perspective on privacy in smart metering," *ACM Trans. Comput. Human Interaction*, vol. 26, no. 1, pp. 1–44, Feb. 2019, doi: 10.1145/3281444.
[40] V. Garg, K. Benton, and L. J. Camp, "The privacy paradox: A Facebook case study," in *Proc. TPRC Conf. Paper*, 2014, pp. 1–40, doi: 10.2139/ssrn.2411672.
[41] M. Levenson, "Nebraska teen who used pills to end pregnancy gets 90 days in jail," *The New York Times*, Jul. 2023. [Online]. Available: https://www.nytimes.com/2023/07/20/us/celeste-burgess-abortion-pill-nebraska.html
[42] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA, USA: Stanford Law Books, 2010.
[43] National Institute of Standards and Technology. "Guide for conducting risk assessments," Sep. 2012. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf
[44] H. Nissenbaum, "Contextual integrity up and down the data food chain," *Theor. Inquiries Law*, vol. 20, no. 1, pp. 221–256, Jan. 2019, doi: 10.1515/til-2019-0008.
[45] B. Allyn, "'The Computer Got It Wrong': How facial recognition led to false arrest of black man," Jun. 2020. [Online]. Available: https://web.archive.org/web/20220823155355/https:/text.npr.org/882683463
[46] S. Karwatzki, M. Trenz, V. K. Tuunainen, and D. Veit, "Adverse consequences of access to individuals' information: An analysis of perceptions and the scope of organisational influence," *Eur. J. Inf. Syst.*, vol. 26, no. 6, pp. 688–715, Aug. 2017, doi: 10.1057/s41303-017-0064-z.
[47] R. Schwinghammer, "Groping in the dark? Exploring customer perception of hidden actions in smart service ecosystems through the lens of agency theory," in *Proc. 57th Hawaii Int. Conf. Syst. Sci.*, 2024, pp. 4268–4277. [Online]. Available: https://hdl.handle.net/10125/106898
[48] M. S. Wogalter, "Communication human information processing (C-HIP) model," in *Handbook of Warnings*, M. S. Wogalter, Ed., Mahwah, NJ, USA: Lawrence Erlbaum, 2006, pp. 51–62.

[49] M. S. Wogalter, D. J. Brems, and E. G. Martin, "Risk perception of common consumer products: Judgments of accident frequency and precautionary intent," *J. Safety Res.*, vol. 24, no. 2, pp. 97–106, 1993, doi: 10.1016/0022-4375(93)90004-7.
[50] M. S. Wogalter and T. Barlow, "Injury severity and likelihood in warnings," *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, vol. 34, no. 8, pp. 580–583, Oct. 1990, doi: 10.1177/154193129003400801.
[51] M. A. Sasse and I. Flechais, "Why do we need it? How do we get it?," in *Security and Usability. Designing Secure Systems that People Can Use*, L. Cranor and S. Garfinkel, Eds. Sebastopol, CA, USA: O'Reilly, 2005, pp. 13–31.
[52] N. Gerber, B. Reinheime, and M. Volkamer, "Investigating people's privacy risk perception," *Proc. Privacy Enhancing Technol.*, vol. 3, pp. 267–288, 2019, doi: 10.2478/popets-2019-0047.
[53] K. R. Laughery, K. P. Vaubel, S. L. Young, J. W. Brelsford, and A. L. Rowe, "Explicitness of consequence information in warnings," *Safety Sci.*, vol. 16, no. 5-6, pp. 597–613, Aug. 1993, doi: 10.1016/0925-7535(93)90025-9.
[54] N. Koester, P. Cichy, D. Antons, and T. O. Salge, "Privacy risk perceptions in the connected car context," in *Proc. Hawaii Int. Conf. Syst. Sci.*, 2021, pp. 4414–4423. [Online]. Available: https://hdl.handle.net/10125/71153
[55] T. Jakobi, M. Von Grafenstein, P. Smieskol, and G. Stevens, "A taxonomy of user-perceived privacy risks to foster accountability of data-based services," *J. Res. Technol.*, vol. 10, Jul. 2022, Art. no. 100029, doi: 10.1016/j.jrt.2022.100029.
[56] P. Mayring, "Qualitative content analysis: Theoretical foundation, basic procedures and software solution," 2014. [Online]. Available: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173
[57] Weizenbaum Institut. *Privacy Icons Project (PIP) Expert Workshop at the Weizenbaum Institute for the Networked Society Research Group 4 - Data as a Means of Payment*. Feb. 2019. [Online]. Available: https://www.weizenbaum-institut.de/media/Permalinks/PIP_Workshop-Short_Report_Engl_7May_FINAL.pdf
[58] M. Meuser and U. Nagel, "The expert interview and changes in knowledge production," in *Interviewing Experts*, A. Bogner, B. Littig, and W. Menz, Eds. London, UK: Palgrave Macmillan, 2009, pp. 17–43.
[59] P. Hanna and S. Mwale, "I'm not with you, yet I am …," in *Collecting Qualitative Data: A Practical Guide to Textual, Media and Virtual Techniques*, V. Braun, Ed., Cambridge, UK: Cambridge Univ. Press, 2017, pp. 235–255.
[60] E. Townsend, E. Nielsen, R. Allister, and S. A. Cassidy, "Key ethical questions for research during the COVID-19 pandemic," *Lancet Psychiatry*, vol. 7, no. 5, pp. 381–383, May 2020, doi: 10.1016/S2215-0366(20)30150-4.
[61] G. Rings and S. Rasinger, "Theoretical approaches," in *The Cambridge Handbook of Intercultural Communication*, G. Rings, A. Ruskin, and S. Rasinger, Eds. Cambridge, UK: Cambridge Univ. Press, 2020, pp. 83–202.
[62] G. D. McCracken, *The Long Interview*. Newbury Park, CA, USA: Sage, 1988.
[63] E. Ivov, "Jitsi meet," Atlassian. [Online]. Available: https://meet.jit.si/
[64] OBS Studio Contributors, "Open broadcaster software | OBS," 2012. [Online]. Available: https://obsproject.com/
[65] S. Drackert, *Die Risiken der Verarbeitung personenbezogener Daten: Eine Untersuchung Zu Den Grundlagen des Datenschutzrechts*. Berlin, Germany: Duncker & Humblot, 2014.
[66] Article 29 Data Protection Working Party, "Opinion 5/2009 on online social networking," 2009. [Online]. Available: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2009/wp163_en.pdf
[67] P. Ohm, "Broken promises of privacy: Responding to the surprising failure of anonymization," *UCLA Law Rev.*, vol. 58, no. 2, pp. 1703–1777, 2009. [Online]. Available: https://www.uclalawreview.org/pdf/57-6-3.pdf
[68] *Safety Aspects Guidelines for Their Inclusion in Standards*, ISO/IEC Guide 51:2014, 2014. [Online]. Available: https://www.iso.org/standard/53940.html
[69] D. Susser, B. Roessler, and H. Nissenbaum, "Technology, autonomy, and manipulation," *Internet Policy Rev.*, vol. 8, no. 2, pp. 1–22, Jun. 2019, doi: 10.14763/2019.2.1410.
[70] S. Barocas and H. Nissenbaum, "Big data's end run around anonymity and consent," in *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, J. Lane, V. Stodden, S. Bender, and H. Nissenbaum, Eds. Cambridge, UK: Cambridge Univ. Press, 2014, pp. 44–76.
[71] M. Favaretto, E. De Clercq, and B. S. Elger, "Big data and discrimination: Perils, promises and solutions. A systematic review," *J. Big Data*, vol. 6, no. 12, pp. 1–27, Feb. 2019, doi: 10.1186/s40537-019-0177-4.
[72] A. Goldfarb and C. Tucker, "Digital Economics," *J. Econ. Literature*, vol. 57, no. 1, pp. 3–43, Mar. 2019, doi: 10.1257/jel.20171452.
[73] Article 29 Data Protection Working Party, "Opinion 03/2013 on purpose limitation," 2013. [Online]. Available: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf
[74] J. P. Choi, J. Doh-Shin, and K. Byung-Cheol, "Privacy and personal data collection with information externalities," *J. Public Econ.*, vol. 173, pp. 113–124, May 2019, doi: 10.1016/j.jpubeco.2019.02.001.
[75] K. R. Laughery and S. Wogalter, "Warnings and risk perception," in *Handbook of Human Factors and Ergonomics*, G. Salvendy, Ed., New York, NY, USA: Wiley, 1997, pp. 1174–1197.
[76] M. Harbach, S. Fahl, and M. Smith, "Who's afraid of which bad wolf? A survey of IT security risk awareness," in *Proc. IEEE 27th Comput. Secur. Found.*, 2014, pp. 97–110, doi: 10.1109/CSF.2014.15.
[77] K. J. Strandburg, "Monitoring, datafication, and consent: Legal approaches to privacy in the Big Data context," in *Privacy, Big Data and the Public Good: Frameworks for Engagement*, J. Lane, V. Stodden, S. Bender, and H. Nissenbaum, Eds. Cambridge, UK: Cambridge Univ. Press, 2014, pp. 44–76.
[78] C. Bravo-Lillo, L. F. Cranor, J. Downs, S. Komanduri, and M. Sleeper, "Improving computer security dialogs," in *Human-Computer Interaction – INTERACT*. Berlin, Germany: Springer, 2011, pp. 18–35, doi: 10.1007/978-3-642-23768-3_2.

[79] M. L. Resnick, "Risk communication for legal, financial, and privacy agreements and mass media," in *Handbook of Warnings*, M. S. Wogalter, Ed., Mahwah, NJ, USA: Lawrence Erlbaum, 2006, pp. 771–782.
[80] A. Rossi and M. Palmirani, "A visualization approach for adaptive consent in the European data protection framework," in *Proc. Conf. E-Democracy Open Government*, 2017, pp. 159–170, doi: 10.1109/CeDEM.2017.23.
[81] H. Habib et al., "Toggles, dollar signs, and triangles: How to (in)effectively convey privacy choices with icons and link texts," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2021, pp. 1–25, doi: 10.1145/3411764.3445387.
[82] European Commission, "Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions A European strategy for data," 2020. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020-DC0066&from=EN
[83] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin, and K. Aberer, "Polisis: Automated analysis and presentation of privacy policies using deep learning," in *Proc. 27th USENIX Secur. Symp.*, 2018, pp. 531–548. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity18/sec18-harkous.pdf

Lukas Seiling received the B.Sc. degree in Psychology from the University of Mannheim, Mannheim, Germany, in 2018. He is currently working toward dual M.Sc. degrees in Cognitive Systems and Human Factors from the University of Potsdam, Potsdam, Germany, and the Technical University Berlin, Berlin, Germany. He is currently a Student Researcher with the Weizenbaum Institute, Berlin, where his research works focus on privacy risk and the design of privacy policies, novel methods for digital social science, and critical perspectives on so-called "artificial intelligence." He has authored or coauthored papers published in *Cognition* and *International Tax Studies*.

Rita Gsenger received dual B.A. degrees in Cultural and Social Anthropology and Philosophy from the University of Vienna, Vienna, Austria, in 2014, the M.A. degree in Philosophy from the University of Innsbruck, Innsbruck, Austria, in 2019, and the M.Sc. degree in Cognitive Science from the University of Vienna in 2021. She is currently a Research Associate with the Weizenbaum Institute, Berlin, Germany, and a Ph.D. candidate with the Institute of Journalism and Communication Studies, Freie University Berlin, Berlin. Her research focuses on the issues of platform regulation, content moderation, evidence-based regulation, and the structural background of disinformation and conspiracy theories, as well as their countermeasures.

Filmona Mulugeta received the B.Sc. degree in Psychology from Technische Universität Braunschweig, Braunschweig, Germany, in 2018, and the M.Sc. degree in Human Factors from Technische Universität Berlin, Berlin, Germany, in 2022. She is currently a Strategy Consultant with Mercedes-Benz Tech Innovation, Berlin. During her studies, she was a part of the research group "Frameworks for Data Markets" at Weizenbaum Institute, Berlin. From 2021 to 2022, she worked on her Master's thesis as part of a research group led by Professor Joachim Meyer at Tel Aviv University, Tel Aviv, Israel. She has authored papers published in 2022 in the *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.

Marte Henningsen received two bachelor's degrees in Computational Engineering and Computer Science from the Leibniz University, Hanover, Germany, in 2019 and 2021, respectively. She is currently working toward the master's degree in Cognitive Science with the University Osnabrück, Osnabrück, Germany, where she is a part of the research group "Ethics and Critical Theories of AI" and mainly works on the impacts of AI systems on society.

Lena Mischau studied law at Ludwig-Maximilians-Universität, Munich, Germany. She completed both German State Examinations in law in 2015 and 2017, respectively. From 2017 to 2021, she was a Research Associate working on the interdisciplinary project related to privacy icons conducted by the research group "Frameworks for Data Markets" at Weizenbaum Institute, Berlin. This article is a follow-up to her coauthored work "Privacy Icons: A Risk-Based Approach to Visualisation of Data Processing," published in the *European Data Protection Law Review* in 2019.

Marie Schirmbeck received the B.Sc. and M.Sc. degrees in Psychology from Humboldt University of Berlin, Berlin, Germany, in 2013 and 2017, respectively, specializing in neurocognitive and engineering psychology. Her work primarily centers on human behavior in various contexts. This article is a follow-up to her coauthored work "Privacy Icons: A Risk-Based Approach to Visualisation of Data Processing," published in the *European Data Protection Law Review* in 2019. Her privacy-related contributions can be found in the white paper "Une Nouvelle Gouvernance pour les Données du XXIe Siècle – Des Standards pour la Circulation et la Protection des Données Personnelles," presented to the French National Assembly in 2019.