

Understanding Online Privacy—A Systematic Review of Privacy Visualizations and Privacy by Design Guidelines

SUSANNE BARTH, DAN IONITA, and PIETER HARTEL, University of Twente, Netherlands

Privacy visualizations help users understand the privacy implications of using an online service. Privacy by Design guidelines provide generally accepted privacy standards for developers of online services. To obtain a comprehensive understanding of online privacy, we review established approaches, distill a unified list of 15 privacy attributes and rank them based on perceived importance by users and privacy experts. We then discuss similarities, explain notable differences, and examine trends in terms of the attributes covered. Finally, we show how our results provide a foundation for user-centric privacy visualizations, inspire best practices for developers, and give structure to privacy policies.

CCS Concepts: • **Security and privacy** → **Usability in security and privacy**; *Privacy protections*; *Social aspects of security and privacy*;

Additional Key Words and Phrases: Privacy attributes, privacy factors, privacy labels, privacy icons, privacy by design

ACM Reference format:

Susanne Barth, Dan Ionita, and Pieter Hartel. 2022. Understanding Online Privacy—A Systematic Review of Privacy Visualizations and Privacy by Design Guidelines. *ACM Comput. Surv.* 55, 3, Article 63 (February 2022), 37 pages.

<https://doi.org/10.1145/3502288>

1 INTRODUCTION

Online services currently handle unprecedented amounts of user-related data [129]. Machine learning algorithms extract value from large amounts of data by recognizing hidden patterns, links, behaviors, trends, identities, and practical knowledge, which has given birth to a “big data economy” [9, 152]. This has opened a “Pandora’s Box” of privacy concerns [113, 141, 151]. But the privacy policies that are meant to address these concerns are often lengthy, legally worded documents written to protect the provider [15, 59]. Even the interactive permission system found on modern smartphones fails to provide a sufficient understanding of the privacy risks involved with using an application [24, 39, 78].

To communicate privacy risks to users in a clear and concise manner, researchers, regulators, and industry have called for a more visual representation of how online services handle personal

This work is supported by the Netherlands Organisation for Scientific Research (NWO) under Grant No.:628.001.011 in collaboration with the Netherlands Organisation for Applied Scientific Research (TNO), Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC), and Centric B.V.

Authors’ address: S. Barth, D. Ionita, and P. Hartel, University of Twente, Drienerlolaan 5, 7522 NB Enschede, Netherlands; emails: {s.barth, d.ionita, pieter.hartel}@utwente.nl.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2022 Association for Computing Machinery.

0360-0300/2022/02-ART63 \$15.00

<https://doi.org/10.1145/3502288>

data [14, 15, 55, 74, 122, 164]. Since 2001, the United States **Federal Trade Commission (FTC)** has been encouraging standardized, tabular privacy policies similar to nutrition labels [13]. The more recent European **General Data Protection Regulation (GDPR)** also suggests using “standardized icons” to provide a meaningful overview of the intended data processing [106]. The Digital Advertising Alliance displays a YourAdChoices button on their ads [47] and the Entertainment Software Rating board has introduced icons indicating whether or not games share personal information with third parties [68]. At the same time, a variety of privacy icons, labels, and notices designed to convey how personal data are handled have been proposed by researchers [60, 62, 66, 77, 123, 145, 147] and industry [63, 115]. However, these visualizations differ with regard to the privacy attributes they cover, as well as their level of detail. Furthermore, the comprehensibility and effectiveness of the visualizations remains questionable as most of them have never been tested with users [122].

Whereas visual representations of privacy attributes are intended for users, **Privacy by Design (PbD)** guidelines are intended for developers. They determine to a significant extent how user privacy is handled. Because developers are not privacy experts, they need clear and unambiguous instructions with regards to how personal data should be handled [36], and they need to know which privacy attributes are considered important by users. While guidelines for what was once referred to as “fair information practice” go as far back as the 1970s [64], technological developments have prompted a renewed interest in developing privacy-aware information systems [127]. However, there is currently no generally accepted PbD standard or best practice. Rather, multiple regulators and industry stakeholders have each elaborated their own PbD principles that, similarly to privacy visualizations, differ significantly in terms of the privacy attributes they consider.

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

2 BACKGROUND

The debate around privacy started in the late 19th century with the launch of the telephone and intensified throughout the “cybernetic revolution” of the 1970s [94]. In his landmark 1967 book, Westin defined privacy as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” [153, p. 7]. Fundamentally, modern privacy is about information [31]. However, the concept kept expanding in both scope and significance with the emergence of the Internet, mass surveillance, terrorism threats [151, 154], and, more recently, with the development of big data and the Web

2.0 [9, 113, 141, 152]. Currently, privacy, and in particular online privacy, remains hard to define [107] or, in the words of Solove, “a concept in disarray” [136]. Smith, Dinev, and Xu [133] notes that historically, privacy was seen as a right, a commodity, a control, or a state. Martin [88, p. 557] sees privacy as a “social contract around what, to whom, and for what purpose” information is gathered or disclosed within a given community and context. Nissenbaum posits that privacy is shaped by social boundaries and norms [99] and, because individuals cannot provide truly informed consent, she suggests articulating context-specific norms that govern the collection and sharing of data online [100]. According to her theory of “contextual integrity,” whether or not an action constitutes a violation of information privacy depends on variables related to the context, the nature of the information, the actors involved and their relationships to the data subject, as well as the terms for collecting and sharing information. Acquisti, Taylor, and Wagman [5] discuss the economic value of privacy and also find that in some situations data sharing can be beneficial for the user, while in other situations it can be damaging. Nevertheless, in his landmark articles, Solove [135, 136] points out that while it is not feasible to arrive at an overarching definition of privacy, the concept can be understood by isolating common “essential” or “core” characteristics. According to Morales-Trujillo et al., to address privacy during software development and to be able to respond to user’s privacy concerns, a conceptual framework is needed that goes beyond data minimization and access control [95].

Solove [137] approached this from a legal perspective by developing a taxonomy of privacy violations pertaining to information collection, information processing, information dissemination, or invasion. From a technical perspective, privacy metrics are often used to compute the efficacy of privacy-enhancing technologies [150], but these are of little use to people without a background in statistics. Martín, del Alamo, and Yelmo [90] highlighted a lack of technical privacy requirements and criticized disagreement between high-level privacy principles. Anwar, Gill, and Beydoun [16] found commonalities between privacy laws and standards but noted differences in nature and scope that require further investigation. Wilson et al. [155] identified 10 categories of data practices by annotating 115 privacy policies. However, Morel and Pardo [96] found that natural language privacy policies required by legislators often cover only a fraction of those categories. They also found significant differences in terms of coverage compared to privacy policies expressed graphically (usually proposed by privacy advocates) or in machine-readable form (usually proposed by academics).

Acquisti et al. [4] saw potential in efforts toward assisting users with online privacy decisions by helping them reflect on their actions before the fact or by “nudging” them toward certain behaviours. But Rossi and Palmirani [122] concluded that existing privacy visualizations vary in terms of the privacy attributes they cover and criticized that the majority were not user tested. They suggest a visual layer summarizing the privacy policy with special focus on the privacy principles of transparency and informed consent but to date, no new system has been developed. Hansen [69] compared privacy pictograms and found most to be of limited practical relevance, noting a lack of international consensus on syntax and semantics. Motti and Caine [97] reviewed icons related to privacy and classified them as either data collection, data transmission, data storage, data sharing, or access control.

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

3 METHOD

Our goal is to distill, validate, and rank a complete list of privacy attributes. The first step toward achieving this was to perform a systematic review to identify privacy visualizations (Section 4.1) and PbD principles (Section 4.2) relevant for online services. We then extracted a list of privacy attributes by coding the results until reaching satisfactory inter-coder reliability and then refining it with practitioners (Section 4.3). Finally, we used online surveys to understand and compare the perceived importance of these privacy attributes to experts and users (Section 4.4). The research methodology behind each of these three steps is described in more detail below.

3.1 Systematic Review

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

We started by searching Scopus using the following queries, selecting papers published between 2001 and 2019.

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

We then followed a systematic review process [132] using a "snowballing" approach [157] described below to iteratively extend the search query and the sample by examining references (Figure 1).

3.1.1 Privacy Visualizations. We read the abstracts and titles of the 2,063 papers retrieved from Scopus and identified 23 that might include an original visual representation of aspects related to privacy or data handling by online services. When scanning these papers, we learned of other terms used to describe privacy visualizations so we assembled these into an extended Scopus and Web of Science query, this time using phrases to reduce the amount of irrelevant results:

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

We read the titles and abstracts of these results and identified 10 more potentially relevant papers. We then read the full text of the 23+10 papers selected thus far and found 17 other relevant papers among their references, which we also read. In total, we were able to find 41 papers containing privacy visualizations. We also learned about a 2016 Workshop on Privacy Indicators but found none of the papers published there satisfied the inclusion criteria described in Section 3.1. To make sure we did not miss anything, we performed several Google searches using all of the keywords we identified and found five more proposals for privacy visualizations coming from **Non-Governmental Organisations (NGOs)** and industry.

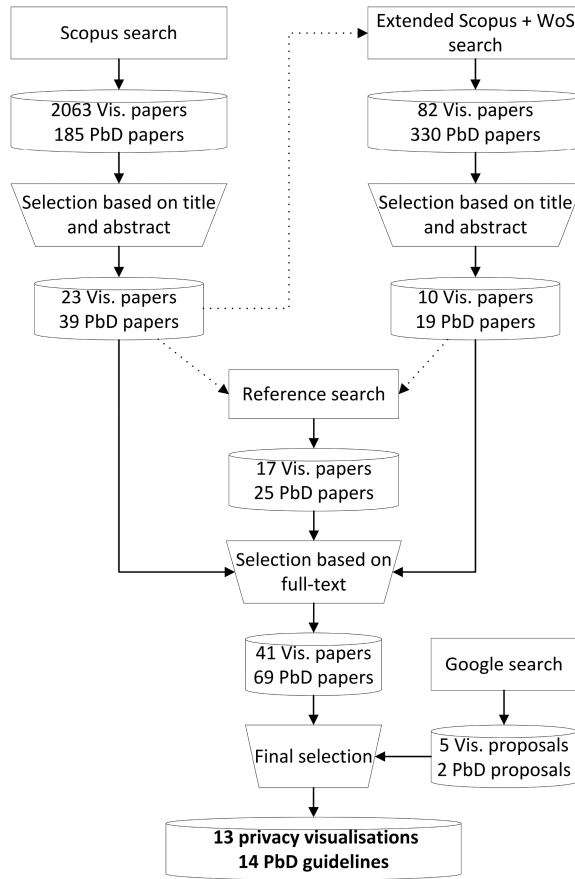


Fig. 1. Systematic review process.

Finally, we analyzed and discussed all of the 41+5 results to select suitable candidates for extracting privacy attributes applicable to online services in general. We therefore excluded papers that were technology specific [57, 65, 76, 131], domain specific [75], or target-group specific [49, 116, 134]. These are marked with an asterisk (*) in our reference list. We also excluded papers that only include an overall rating [18, 46, 71, 138, 163], as well as other papers where no individual attributes could be distinguished [56, 83, 118]. Finally, we excluded papers that evaluate and classify existing visualizations [69, 96, 144, 164], because the visualizations they cover were already in our sample.

After removing duplicates, we ended up with a final sample of 13 privacy visualizations that we discuss in detail in Section 4.1 and that served as input for our coding process. Of these 13, 7 come from academia, 5 from industry, and 1 from government.

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

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

We read the titles and abstracts of these results and identified 18 more potentially relevant papers. We then read the full text of the 39+18 papers selected thus far and found 25 other relevant papers among their references, which were also read. In total, 69 papers containing high-level PbD principles were found. We also ran a Google search based on the original Scopus query and found two other proposals for PbD principles coming from industry.

We analyzed and discussed all of the 69+2 results to select which are suitable for extracting generally applicable privacy attributes. We therefore excluded papers that discuss technology-specific principles [2, 61, 109–112, 114, 130, 148] as well as papers that translate generic PbD principles to specific domains [21, 27, 28, 35, 38, 48, 82, 117, 124, 146, 149]. These are marked with a dagger (†) in the Reference list. We excluded papers that discuss and compare existing PbD principles [22, 70, 85, 120, 121, 128] as well as those that refine or operationalize PbD principles [6, 10, 19, 26, 37, 41–43, 50, 52, 91, 92, 139, 140, 143], but added the PbD principles they reference to our sample.

After removing duplicates, we ended up with a final sample of 14 PbD guidelines that we discuss in detail in Section 4.2 and that served as input for our coding process. Of these 14, 2 come from academia, 5 from industry, and 7 from government.

3.2 Coding

To analyze the results of the systematic review, we followed an iterative coding process.

First, the second author of this article analyzed the privacy visualizations and PbD guidelines selected during the systematic review. The content was divided into passages, and each passage was coded with one or more terms related to the handling of personal data. This resulted in an initial list of 13 privacy attributes.

Second, we discussed the initial list of privacy attributes with two information security professionals from a large software solutions provider in a 1-hour unstructured interview. Both security professionals deal with information privacy on a daily basis. As a result of the interviews, two attributes were split-up and the definitions of the attributes were clarified.

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

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

3.3 Online Survey

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

The survey, approved by the ethical committee of The University of Twente, collected demographic data about gender, education, occupation, nationality, and the type and frequency of online service usage. We asked the subjects how important on a scale from 0 (not at all important)

to 10 (extremely important) they considered each of the 15 privacy attributes. The full phrasing is offered as the supplementary material. Since we want to obtain an overall ranking, we did not select a specific scenario. To assess the sensitivity of our findings, we asked participants whether or not they would rate these attributes differently for different types of services. Finally, in open questions, we asked if any of the descriptions were ambiguous and if they felt any attributes were missing.

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

4 RESULTS

4.1 Privacy Visualizations

Privacy visualizations are visual representations designed to communicate aspects related to the handling of personal data to users of online services. In this section, we briefly describe the 13 privacy visualizations selected in chronological order and discuss the privacy attributes they cover. The complete icon sets are provided as supplementary material.

4.1.1 Mehldau's Data-privacy Declarations. To the best of our knowledge, designer Martin Mehldau was the first to propose an iconset to communicate the privacy aspects of an online service, in 2007. His list of “data-privacy declarations” contained 30 icons grouped into four categories that could be used to represent how data are used, stored, shared, or deleted. Because of the large number of icons, we do not show them here.¹ Some examples per category:

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

4.1.2 KnowPrivacy's Policy Coding Methodology. In 2009, the KnowPrivacy research project² proposed set of “tags” used for coding privacy policies. Each tag referred to a *type of user data*, a *general data practice*, or a *data sharing* agreement and consisted of an icon and a description, as shown in Figure 2. For a given privacy policy, each tag could be in one of three states: YES, NO, or UNCLEAR [66] to visually indicate what data are collected, how they are used, and who they are shared with. They coded 50 privacy policies and compared the results with consumer expectations, finding a “large level of ignorance on the part of users about how data is collected” [79].

¹The full list of icons is available under a CC-BY license from: <https://netzpolitik.org/wp-upload/data-privacy-icons-v01.pdf>.

²<http://knowprivacy.org>.














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

Fig. 2. KnowPrivacy codes.

4.1.3 CyLab’s Privacy Nutrition Label. Developed by Carnegie Mellon’s CyLab Usable Privacy and Security laboratory in 2009, the privacy nutrition label [77] takes a tabular approach to represent how personal data are handled by an online service provider (see Figure 3). Each row corresponds to a data item (e.g., location, health information, etc.), and each column corresponds to a way in which each item is used (e.g., marketing, profiling, sharing with other companies, etc.). Each cell in the resulting matrix gives a visual indication with regard to each data item–usage pair as follows:

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

The rows and columns are fixed so that two policies can be compared side by side. There are a total of 10 data items and seven ways in which these can be used. The possible usages are as follows: (1) *provide service and maintain site*, (2) *research and development*, (3) *marketing*, (4) *telemarketing*, (5) *profiling*, (6) *sharing with other companies*, and (7) *sharing on public forums*.

4.1.4 Mozilla’s Privacy Icons. In 2010, Aza Raskin from Mozilla proposed a set of icons that could be attached to existing privacy policies to provide a visual summary of the most important privacy issues: *retention period*, *third-party use*, *ad networks*, and *law enforcement*. The icon designs to represent these attributes have been the subject of multiple iterations, the latest are show in Figure 4. The project has since been abandoned but the icons are still present on Mozilla’s Wiki.³

³https://wiki.mozilla.org/Privacy_Icons.

Bell Group

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

Fig. 3. Privacy nutrition label example.

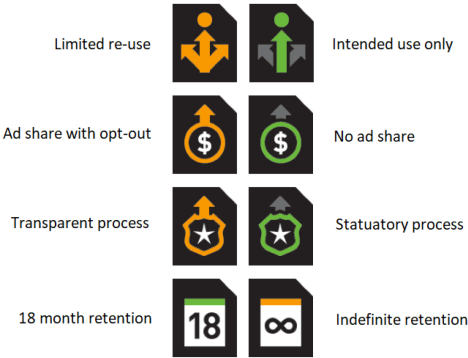


Fig. 4. Mozilla privacy icons v2.



Fig. 5. Highest rated icons from the PrimeLife project.



Fig. 6. Icons of the TrustArc short notice.

4.1.5 *The PrimeLife Project.* Also in 2010, the EU-funded PrimeLife project⁴ published several sets of icons: a general set and other sets for specific domains such as social media [60]. The icons were designed to be aligned with European privacy laws. The initial proposal contained 30 icons representing three types of privacy concepts: *data types* (i.e., personal, sensitive, payment, or medical data), *data purpose* (i.e., legal obligation, shipping, tracking, or profiling), and *data processing* (storage, deletion, pseudonymization, anonymization, disclosure, and collection). For

⁴<http://primelife.ercim.eu>.



Fig. 7. The privacy wheel.

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

Fig. 8. Privacy icons from GDPR draft.

social networks, PrimeLife added icons for groups of recipients (friends, friends of friends, selected individuals, and public). They performed user studies to compare different designs and found that icons should be as simple as possible and culturally neutral, and their number held to a minimum. [74]. Figure 5 shows the icons rated highest during their evaluation.

4.1.6 TrustArc’s Privacy Short Notice. In 2011, TrustArc (the developers of the TRUSTe privacy certification standard) proposed an icon-based “privacy short notice” aimed at providing a simplified summary of privacy policies. After analyzing previous approaches, they concluded that such a short notice should focus on the data practices and uses that are invisible to users, namely *secondary use* (none, customization, or profiling), *sharing* (none, affiliates, or unrelated), *third-party tracking*, and *data retention* (none, limited, or indefinite) [115]. Therefore, their visualization only includes four icons (see Figure 6), accompanied by textual descriptions.

4.1.7 Privacy Wheel. Based on a survey that showed users prefer general and less legally detailed information about data handling practices, van den Berg and van der Hof [145] developed the privacy wheel. Taking the privacy principles of the **Organisation for Economic Cooperation and Development (OECD)** as a starting point for their visualization, the wheel (see Figure 7) covers eight core concepts of privacy related information: (1) *collection*, (2) *data quality*, (3) *purpose*, (4) *limited use*, (5) *security*, (6) *consent*, (7) *third parties*, and (8) *accountability*. The spokes of the wheel are clickable, providing two layers of increasingly detailed information. Furthermore, some spokes provide an interactive mechanism for updating opt-in/opt-out preferences. The service provider has the obligation to apply those changes in the system.

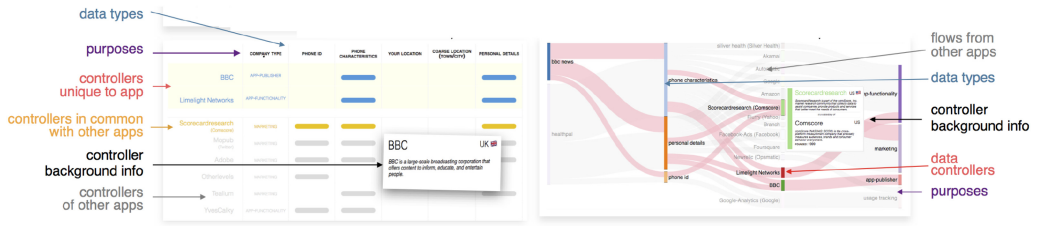


Fig. 9. Example DCIs label.

4.1.8 GDPR’s Draft Privacy Icons. Article 12 of the European GDPR mandates that “The information ... may be provided in combination with standardized icons to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing.” [106]. In addition, it specifies that the icons should be machine readable. The final version of the GDPR does not prescribe specific icons or specific attributes that need to be represented but does empower the European Commission to determine these at a later time. However, an earlier draft of the GDPR did explicitly describe six icons shown in Figure 8. For each icon, a given application may score a checkmark or an X.

4.1.9 DCIs. Developed in 2017 by researchers from the University of Oxford and Cambridge [147], the **Data Controller Indicators (DCIs)** provide information on the kinds of data that are sent by an app to various parties while considering the background information of those parties and the purposes behind data usage. Unlike the other visualizations we reviewed, DCIs labels are automatically generated and therefore can be easily scaled to a large number of services. Testing different versions of the visualization with users revealed a preference for Personalized DCIs (see Figure 9) that provide a differential risk assessment of data controllers by comparing the dataflows of multiple apps.

4.1.10 Fox et al.’s GDPR Compliant Label. In 2018, Fox et al. [62] started developing a privacy label that is compliant to the requirements mandated in the GDPR. Their label is based on the CyLab’s privacy nutrition label. In an iterative process, the authors developed an icon- and a text-based label and tested them in the context of an e-commerce website, revealing users’ preference for the icons. Consequently, an icon-based label was further developed, covering 12 privacy attributes as shown in Figure 10: (1) *information about data controller*, (2) *data processing purposes*, (3) *recipients of personal data*, (4) *transfer to third countries*, (5) *retention*, (6) *rights of data subject*, (7) *consent*, (8) *right to complain*, (9) *disclosure*, (10) *automated decision-making*, (11) *details of data protection officer*, and (12) *further data processing*. Next, the authors aim to test the label with and without ON/OFF toggles to indicate consent.

4.1.11 CLEVER°FRANKE’s Privacy Label. In 2019, SensorLabs, a Dutch non-profit initiative by UX design firm CLEVER°FRANKE, published a highly simplified privacy label [63]. The label is designed for online services as well as physical devices, such as vending machines, card scanners, and even storefronts. To come up with the label, they reviewed the literature on conceptualizing and extracted three essential aspects of privacy, namely (1) *collection*, (2) *purpose*, and (3) *control*. Each of these aspects is measured using five yes/no questions based on the Rathenau Institute’s overview of ethical and societal issues related to digitization [80]. Each “yes” answer achieves 1 point, up to a maximum of 15 points. The final score determines what label the entity receives. Each label consists of two elements: an A-to-F category that also determines the color (A is green, F is red, everything in between is shades of orange), and a visual representation of the score on



Fig. 10. Example of Fox et al.'s GDPR compliant label.



Fig. 11. Five example labels from CLEVER²FRANKE.

each of the three aspects. Figure 11 shows some example labels. The circle around the letter is divided into three parts corresponding to collection, purpose, and control. Each part consists of five layers corresponding to the five questions for each aspect.

4.1.12 DaPIS. The Data Protection Icon Set, developed by Rossi and Palmirani in 2019 [123], is based on PrOnto, a computational ontology of the GDPR. The machine-readable layers provide interpretable information from legal documents, whereas the human-centered layer adds visual accessible icon design. As seen in Figure 12, DaPIS covers: (1) *data*, e.g., personal; (2) *agents' roles* e.g., data subject or controller; (3) *processing operations*, e.g., anonymization or profiling; (4) *data subject's rights*, e.g., access or erasure; (5) *processing purposes*, e.g., research or marketing; and (6) *legal bases for processing*, e.g., consent or legitimate interest. The authors emphasize that DaPIS is not designed to be a standardized European icon set but provides a foundation for the implementation of GDPR's icons and is still under development.

4.1.13 Privacy Label (privacylabel.org). In 2020, a Dutch consortium of privacy related companies and non-profit organizations, launched privacylabel.org. In combination with graphical icons, their tabular label provides information on seven core themes: (1) *data collection*, (2) *purpose*, (3) *data sharing*, (4) *location*, (5) *duration*, (6) *legal grounds*, and (7) *take action*. See Figure 13 for

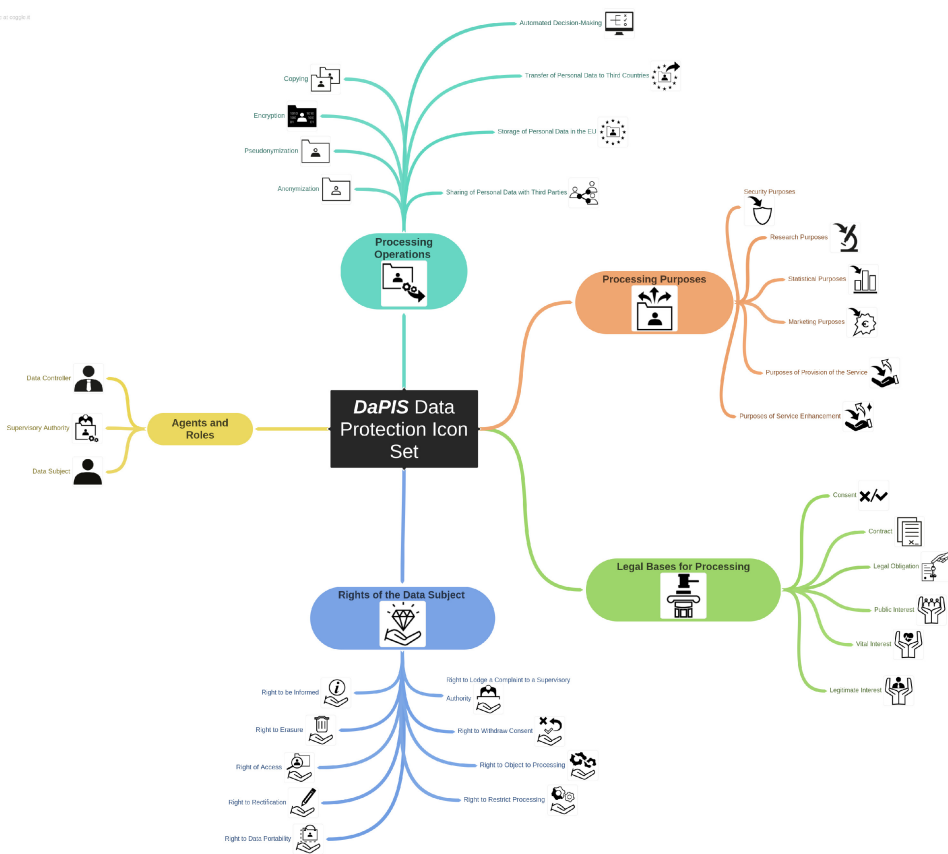


Fig. 12. The data protection icon set (DaPIS).

an example. The title of each theme is clickable, providing short explanations in relation with the GDPR regulations. For more information, a “learn more” option is provided, directing the user to the Privacy Label website. Furthermore, each core theme contains several sub-themes (referred to as “ingredients”) that are, in turn, clickable and provide further information tailored to the data practices of the online service (e.g., the reason behind data aggregation).

4.2 Privacy by Design Guidelines

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

4.2.1 The Australian Privacy Principles. The **Australian Privacy Principles (APPs)** were first added to the Australian Privacy Act in 2001. The APPs apply to the private sector and most government entities in Australia. They are technology neutral, and can be tailored to the needs of individual organizations. In 2014, the original list of 10 principles was extended to 13 [102]:

- *Open and transparent management of personal information:* Manage personal data in an open and transparent way, including a clear and up-to-date privacy policy.

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

Fig. 13. Example of a label generated by privacylabel.org.

- *Anonymity and pseudonymity*: Provide individuals the option of not identifying themselves.
- *Collection of solicited personal information*: Conditions for collecting personal or sensitive data when needed and allowed.
- *Dealing with unsolicited personal information*: Avoid gathering unsolicited personal data.
- *Notification of the collection of personal information*: Provide information about data collection.
- *Use or disclosure of personal information*: Conditions for usage or disclosure of personal data.
- *Direct marketing*: Restrict use, disclosure of personal data for direct marketing purposes.
- *Cross-border disclosure of personal information*: Conditions for personal data protection before disclosure overseas.
- *Adoption, use or disclosure of government related identifiers*: Conditions for government related identifier adoption, or the disclosure of it.
- *Quality of personal information*: Ensure personal data collected, used, or disclosed is accurate, up-to-date, and complete.
- *Security of personal information*: Protect personal data and remove it when needed.
- *Access to personal information*: Conditions for providing access to personal data.
- *Correction of personal information*: Obligations for amendment of personal data.

4.2.2 CSA's Model Code for the Protection of Personal Information. First published in 1996, the **Canadian Standards Association (CSA)** has reaffirmed the Model Code for the Protection of Personal Information in 2001. The standard is focused around privacy rights and individual control over the use and exchange of personal information. Eventually, the 10 principles developed by the CSA have been incorporated into Canadian law. The following principles form the basis of the Model Code for the Protection of Personal Information [67]:

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

4.2.3 *APEC's Privacy Framework*. Published in 2005, **Asia-Pacific Economic Cooperation (APEC)** developed a principle-based privacy framework. Inspired by OECD guidelines, this framework aims at developing information privacy protections and to warrant the free information flow in the Asia Pacific region. The privacy framework includes the following privacy principles [45]:

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

4.2.4 *The Global Privacy Standard*. The **Global Privacy Standard (GPS)**, was published in 2006, at the 28th International Data Protection and Privacy Commissioners Conference. Its purpose was to reinforce the mandate of data protection authorities by drafting “fundamental and universal privacy concepts” [32], namely:

- *Consent*: Consent for collection, use or disclosure of personal information, and ability to withdraw consent.
- *Accountability*: Communicate all privacy policies and procedures and seek equivalent privacy protection from third parties.
- *Purposes*: Specify and communicate the purpose for collecting, using, retaining and disclosing personal information.
- *Collection Limitation and Data Minimization*: Collection is fair, lawful and limited to specified purposes; data minimization and anonymization or pseudonymization should be applied.
- *Use, Retention, and Disclosure Limitation*: Limit use, retention, and disclosure of personal information to specified purposes, except when required by law.
- *Accuracy*: Accurate, complete, up-to-date personal information as per the specified purposes.
- *Security*: Ensure security of personal information throughout its lifecycle as per recognized international standards.
- *Openness*: Make information about policies and practices related to personal information readily available.
- *Access*: Provide access to personal information, its uses, and allow to challenge its completeness or have it amended.
- *Compliance*: Monitor, evaluate, and verify compliance with privacy policies and procedures.

4.2.5 ISTPA's Privacy Framework. In 2007, triggered by considerable changes in information privacy since 2002, as well as huge variations in the language and content of existing privacy frameworks, the **International Security, Trust and Privacy Alliance (ISTPA)** performed a structured review of existing privacy regulations and standards and extracted a set of key principles. They supplemented the list with three additional principles, resulting in a working set of 11 privacy principles [126]:

- *Notice*: Provision of an overarching privacy policy.
- *Consent*: Opt-in/opt-out, or implied affirmative process.
- *Collection Limitation*: Minimal data collection and related to purposes.
- *Use Limitation*: Usage and retention of personal data for specified purposes only.
- *Disclosure*: Release, transfer, access or re-use of data with consent of data subject only.
- *Access and Correction*: Ability to access and amend personal data.
- *Security/Safeguards*: Confidentiality, availability and integrity of personal data.
- *Data Quality*: Adequacy, up-to-dateness, minimization or elimination of personal data in relation to purposes.
- *Enforcement*: Assurance of compliance with privacy policy and ability to challenge this.
- *Openness*: Availability of privacy policy.
- *Anonymity*: Prevention of identification.
- *Data Flow*: Communication of data across geo-political jurisdictions.
- *Sensitivity*: Specification of data that need special security controls.

4.2.6 The Generally Accepted Privacy Principles. In 2009, the American Institute of Certified Public Accountants and the Canadian Institute of Chartered Accountants published **Generally Accepted Privacy Principles (GAPP)** [7]. It was intended as a global privacy framework aimed at helping accountants develop their own privacy program. GAPP is supported by 70 objectives grouped under 10 core principles:

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

4.2.7 Cavoukian's 7 Foundational Principles. Also in 2009, Ann Cavoukian [33, 34], the Privacy Commissioner of Ontario combined Langheinrich's Principles of Privacy-Aware Ubiquitous Systems [82] with those of the GPS [32] into a set of high-level design principles for privacy-aware software that were later adopted by Deloitte [30]:

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

4.2.8 ISO29100 Privacy Framework. In 2011, the ISO/IEC Information Technology Task Force published its own privacy framework specifying a common privacy terminology while defining actors and roles involved in the processing of **Personally Identifiable Information (PII)** [1]. Revised in 2017, the standard defines its own set of privacy safeguarding considerations, namely:

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

4.2.9 OECD's Privacy Principles. Based on their 1980 Fair Information Practices aimed at the trans-border flow of information, the OECD published a revised set of privacy principles in 2013 that integrated the recent work on privacy law enforcement cooperation, resulting in the following principles [101]:

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

4.2.10 *Hoepman's Privacy Design Strategies*. First published in 2014, Hoepman's *Little Blue Book* [72] outlines strategies to make PbD more concrete and applicable in practice. The book translates legal norms and best-practices surrounding personal data into the following design requirements:

- *Minimise*: Keep the amount of personal information processed to a minimum.
- *Hide*: Hide any personal information that is processed from plain view.
- *Separate*: Process personal information in a distributed fashion whenever possible.
- *Aggregate*: Process personal information at the highest aggregation and with the least detail.
- *Inform*: Inform data subjects adequately whenever personal information is processed.
- *Control*: Provide data subjects with agency over the processing of their personal information.
- *Enforce*: Have a privacy policy compatible with legal requirements in place and enforce it.
- *Demonstrate*: Demonstrate compliance with privacy policy and legal requirements.

4.2.11 *OASIS Privacy Management Reference Model*. The **Privacy Management Reference Model and Methodology (PMRM)** was developed and published in 2016 by the **Organization for the Advancement of Structured Information Standards (OASIS)**, a non-profit organization committed to privacy and personal data protection. Derived from international legislation and regulations, the PMRM provides a set of 14 privacy principles [51]:

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

4.2.12 *The Privacy Company's PbD Framework*. In 2018, the Privacy Company published a data protection by design framework aimed at developers [29]. It translates the requirements of the European GDPR into the following guidelines:

- *Anonymization*: Anonymize and aggregate.
- *Data minimization*: Gather only necessary data and delete unnecessary data immediately.

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

4.2.13 GDPR Art. 5. Launched in 2018, the European GDPR regulates data privacy laws across Europe and replaced the Data Protection Directive 95/46/EC. All organizations that target or collect data from people within EU must comply with the GDPR. Article 5 of the GDPR covers the following seven data protection principles relating to the processing of personal data [106]:

- *Lawfulness, fairness and transparency*: Lawful, fair and transparent processing.
- *Purpose limitation*: Specification of legitimate purposes for data processing.
- *Data minimization*: Collection and processing restricted to what is absolutely necessary.
- *Accuracy*: Data kept accurate and up-to-date.
- *Storage limitation*: Storage only as long as necessary for purpose fulfillment.
- *Integrity and confidentiality*: Appropriate security, integrity, and confidentiality.
- *Accountability*: Responsibility for compliance with these principles.

4.2.14 The Personal Information Protection and Electronic Documents Act. Under the authority of the Office of the Privacy Commissioner of Canada, the **Personal Information Protection and Electronic Documents Act (PIPEDA)** was revised in 2019. PIPEDA provides 10 fair information principles that serve as the groundwork for the collection, use, disclosure of, and access to personal data handled by the private sector [103]:

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

4.3 Unified List of Privacy Attributes

By means of an open-coding procedure, we distilled an initial list of 13 privacy attributes from the privacy visualizations and PbD guidelines we reviewed: accountability, anonymization, collection, control, correctness, disclosure, functionality, purpose, retention, sale, security, sharing, and transparency. After discussing this list with two practitioners (see Section 3.2), we added pseudonymization and the right to be forgotten before adding simple definitions to each of the privacy attributes.

Finally, we iteratively refined the definitions during three rounds of coding, arriving at the following unified list of privacy attributes, ordered alphabetically:

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

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

- High level data aggregation is part of anonymization.

Collection = Which data are collected? *e.g., IP address, phone number, credit card information, and so on*

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

Control = Must the data subject provide consent for collection and processing of their data and to what extent is the data subject able to opt-out of data collection or processing?

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

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

- Correctness has to do with the ability of the service provider and/or is the user able to *fix* incorrect data after the data were collected?
- Correctness goes a step further than control: If data are already disclosed, is the user able to correct data about him or herself that is not (or no longer) valid?

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

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

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

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

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

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

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

Retention = How long is the collected data stored?

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

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

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

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

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

Sharing = Does any of the collected data leave the ownership of the service provider? *e.g., other companies, advertisers, research institutions, and so on*

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

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

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

Table 1 shows which attributes were covered by privacy visualization or PbD guidelines. Notably, most privacy visualizations and PbD guidelines cover issues regarding collection and purpose. However, data sharing is only covered by half of the PbD guidelines. Furthermore, while all PbD guidelines make statements about security and transparency requirements, only half of the privacy visualizations we reviewed communicate these aspects to users. Accountability and correctness are also mentioned frequently in PbD principles but were rarely covered by privacy visualizations. Functionality was only found in Cavoukian's PbD guidelines and CLEVER°FRANKE's privacy label, and sale of data is only covered by two privacy visualizations and zero PbD guidelines. The similarities and differences are discussed in detail in Section 5. For a detailed overview of how each proposal defines each of the attribute, we refer the reader to our Supplementary Material.

4.4 Perceived Importance of Privacy Attributes

As part of the online survey described in Section 3.3, 385 users and 100 privacy experts ranked the importance of the privacy attributes as described in Section 4.3. Figure 14 shows the mean importance of each attribute for the users and the privacy experts as well as the 95% confidence intervals. Notably:

- Both users and privacy experts in our study agree that collection, sharing, and sale are the most important privacy attributes.
- Privacy experts assign up to about 10% *more* importance than users to most attributes. However, the same experts assign anonymization and the right to be forgotten with up to 10% *less* importance than users.
- The mean scores of users and privacy experts differed most for retention (+1.031), $t(473) = 3.547$, $p = 0.000$, purpose (+0.934), $t(223) = 5.049$, $p = 0.000$, and sale (+0.822), $t(221) = 4.538$, $p = 0.000$.

In our sample, 59% of privacy experts and 49% of users indicated that they would rate the attributes differently for different types of services. This is in line with similar findings indicating that privacy concerns are dependent on the context and the type of service [89, 99, 100, 113, 159].

ACM Computing Surveys, Vol. 55, No. 3, Article 63. Publication date: February 2022.

[illegible]

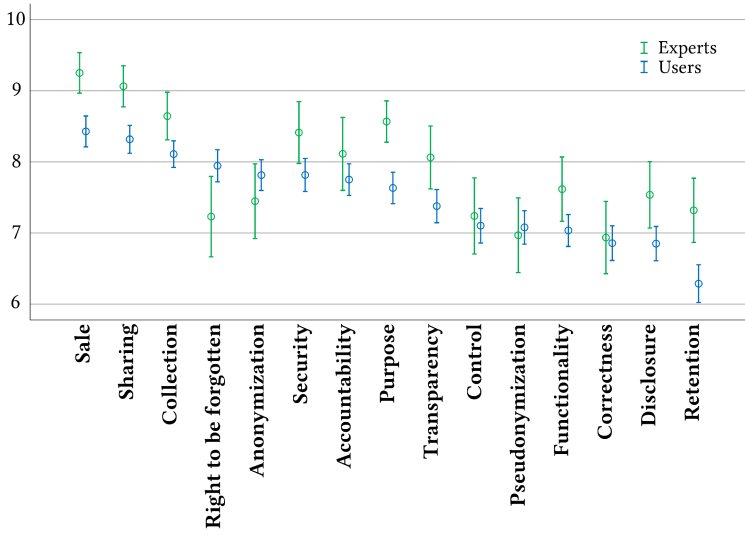


Fig. 14. Mean importance (0-to-10) and confidence interval of privacy attributes, sorted by mean importance to users.

Previous research suggests that privacy concerns are influenced by demographic factors [23, 165]. To investigate whether men and women felt differently about their privacy, we ran an independent sample t -test for all 15 privacy attributes. No significant differences were found, which is consistent with the results of a recent meta-study [142]. Since age is often found to be associated with privacy expectations [23], we ran an ANCOVA to control for the age of the respondents in assessing the differences between the mean scores of the users and the privacy experts. The only significant difference we found was for the right to be forgotten ($p = 0.005$), but the adjusted means were almost the same as the unadjusted means. Therefore we conclude that age is not a confounding variable. The vast majority of the respondents were European nationals (41% Dutch, 23% German, 18% British). Since all Europeans fall under the same privacy regime, controlling for nationality was deemed unnecessary.

5 DISCUSSION

Our literature review (Table 1) revealed notable differences between privacy visualizations and PbD guidelines in terms of the privacy attributes they cover. And, on average, PbD guidelines cover more attributes than visualizations (8.6 vs. 5.6 attributes per proposal). This result is not surprising if we consider that privacy visualizations are mostly designed to provide simple, user-friendly information about the handling of personal data [93, 105].

Additionally, experts and users rated some attributes differently in the survey of Section 4.4. And, overall, privacy experts assigned a higher importance to most attributes. This is to be expected, because as privacy officers, they are not only concerned with privacy as users, but also professionally.

Sale and *sharing* were rated as the most important attributes by most users and privacy experts in our sample. However, while icons related to the *sharing* of data were included in all but one of the privacy visualizations, only half of the PbD principles dealt with this issue. Studies find that willingness to exchange personal data is strongly mitigated by secondary use [89, 137] so it makes sense that almost all of the privacy visualizations we reviewed describe data sharing, since

they are aimed at users. Sale, the attribute consistently ranked as most important in our survey, is covered by just two privacy visualizations and zero PbD guidelines. This is evidence of a growing discrepancy: While the sale of personal data remains an intrinsic part of the business model for online service providers [87, 125] it is one of the major concerns of users [81].

Collection and *purpose* are arguably the most fundamental privacy attributes, because they describe which data are to be collected and why. The privacy experts we surveyed consider both collection and purpose to be of very high importance (closely following sale and sharing). Other studies confirm this observation [3]. Users in our sample, however, rate purpose as less important. We speculate this is because users consider certain types of data as sensitive regardless of purpose [20]. Nevertheless, collection and purpose were the most frequently occurring attributes in both privacy visualizations and PbD guidelines. Therefore, they appear to be the most important attributes to consider when discussing online privacy.

Transparency was mentioned in all PbD guidelines but only half of the privacy visualizations, even though users value insights in data handling practices [144]. We speculate this is because privacy visualizations are themselves a tool for transparency. Nevertheless, complete transparency can only be achieved by having access to the source code or raw data streams [84].

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

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

Retention is ranked significantly higher by privacy experts compared to users and also covered by most PbD guidelines and privacy visualizations. *The right to be forgotten*, however, was perceived as more important by users and is rarely mentioned in the privacy visualizations or PbD guidelines we reviewed. The right to be forgotten and retention both relate to the ability of an organization to delete privacy sensitive data. Retention has always been a technical consideration, but the right to be forgotten is a relatively new, user-driven initiative. This is supported by the fact that in our literature review, we only found one mention of it before 2011. However, managing legacy data sources in a GDPR-compliant manner is a major challenge [112]. Knowing how hard it is to completely remove data from all sources might cause privacy experts to rate the importance of the right to be forgotten lower than retention. However, users are likely more interested in the benefits such a right would provide rather than the technical constraints.

Anonymization was ranked as the fifth most important attribute by users and the eighth most important attribute by privacy experts, but received surprisingly little attention in the literature reviewed here. Anonymization is technically challenging [95] and privacy experts know this. Because true anonymization is seldom achievable [108, 158], various degrees of pseudonymity are implemented instead. Although the information security practitioners in our initial focus group felt that *pseudonymization* should be differentiated from anonymization, several privacy experts in our survey indicated that the two attributes are difficult to distinguish. We speculate users are also not familiar with this distinction and ranked pseudonymization as less important, because

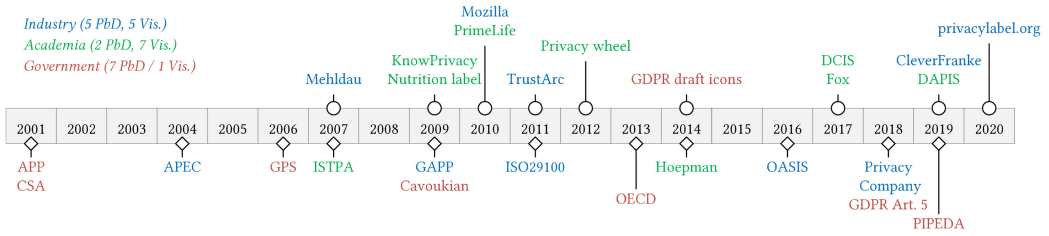


Fig. 15. Publication timeline of privacy visualizations and PbD guidelines.

it implies less protection. Nevertheless, taking steps to remove personal identifiers from user data is of interest to users, which also implies this should be given more careful consideration by developers. However, from a practical perspective, pseudonymization can be viewed as partial or imperfect anonymization.

Control and *correctness* were ranked relatively low by both users and privacy experts but were often encountered in PbD guidelines. Furthermore, correctness was represented in 30% of visualizations. Online services increasingly gather and aggregate user data to glean insights into habits, trends or behaviors not directly related to the actual exchange of the product or service [9, 87, 152], but privacy controls are widely perceived as overly complex by users [25, 119]. The resulting difficulty in managing personal data results in privacy fatigue: a sense of not being in control of the collection and sharing of data online [40, 73]. This weakens the perceived utility and therefore importance of privacy settings and controls. Nevertheless, such mechanisms enhance privacy both proactively (preventing unauthorized collection or collection of incorrect data) and reactively (consent withdrawal and correction of previously collected data). Therefore, providing control over data collection and maintaining correctness of user data is an inherent part of online privacy [87].

5.1 Trends

Although we reviewed PbD guidelines published or updated after 2001 (see Figure 15), our initial search returned many older PbD guidelines. The FTC Fair Information Practice was the first set of PbD principles, forming the foundation for many of the newer principles and legislation. In 1990, the **United Nations (UN)** published similar guidelines and in 1995, the **European Union (EU)** introduced its first Data Protection Directive. Throughout the first decades of the 21st century, the publication rate of PbD guidelines slowly increased and after 2009 we saw an increase in domain- or technology-specific PbD guidelines. Since most of the PbD guidelines we found are either regulation or industry standards, we conclude that PbD by design has made its way into practice.

However, all of the privacy visualizations we found were published after 2007, with the majority being published by academics after 2012. This coincides with an increase in privacy awareness. Although the need for communicating online privacy is not a new discussion [93], research into empowering users to make informed disclosure decisions has recently started to gather steam [55, 74, 105, 122]. We are starting to see industry initiatives as well. However, despite the fact that both the European GDPR and the U.S. FTC recommend standardized privacy labels, no official standard has yet been defined.

Disclosure, *correctness*, *accountability*, and *the right to be forgotten* are increasingly common in recent privacy visualizations. This trend likely reflects increasing concerns regarding safe harbor [44] and data breaches [54]. Even though correctness and accountability are covered by many PbD guidelines, disclosure is not covered by recent initiatives such as PIPEDA, the Privacy Company, and privacylabel.org.

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

5.2 Limitations

Because the entry point of the literature search was Scopus, it is possible that not all relevant proposals from industry were considered. We mitigated this by performing auxiliary Google and Web of Science searches. Furthermore, even though we ran several searches using nine synonyms for principles and 14 synonyms for visualizations, important keywords may have been missed. We do believe, however, that our literature sample of 27 proposals is sufficient to reach saturation in terms of privacy attributes. This is supported by the fact that each privacy attribute was encountered in at least two documents and that over 93% of privacy experts and users we surveyed indicated the unified list was complete and unambiguous.

Some of the documents selected for our systematic review were ambiguous and many differed in terms of granularity and scope. Therefore, multiple attributes were sometimes attached to the same principle or visualization and multiple principles or visualizations sometimes corresponded to a single privacy attribute. Nevertheless, after three rounds of coding we reached almost perfect agreement between coders. This indicates an inherent overlap between the attributes that is to be expected, because they are inter-dependent and refer to the same overarching concept. Nevertheless, while the attributes on our list could be grouped, broken up, or renamed for practical applications, the list itself is complete and understandable.

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

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

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

5.3 Practical Recommendations

5.3.1 Privacy Visualizations Should Be Legally Mandated. Except for CLEVER°FRANKE's, DAPIS and privacylabel.org, which are currently under development, all of the other privacy visualization projects have been abandoned. We speculate that adopting such labels—and more importantly, getting a good score—provides a non-functional benefit to the user but comes at great costs for the provider, as is often the case with safety and security. Indeed, third-party privacy seals are not correlated with trustworthiness [53] and crowd-sourcing efforts such as **Terms of Service; Didn't Read (TOS:DR)** have so far been unsuccessful. Providers should therefore supply an understandable summary of their privacy policies themselves [151]. However, since similar en-

deavors such as the EU energy label, movie ratings, and even seatbelts had to become mandatory before they were adopted, privacy visualizations will only become wide-spread if they are legally mandated.

5.3.2 Privacy Visualizations Should Go Beyond Data Collection and Processing. We find that most privacy labels align with Nissenbaum [100] and Martin and Shilton [89] in that they primarily communicate what information is collected, how this information is shared and for what purpose. However, our ranking suggests that sale of data must also be made explicit. Furthermore, although most current visualizations do not include an indication of the level of security and accountability, this is important to both privacy experts and users and actually mandated by the GDPR [106]. Trustworthy online data exchange relies on obtaining truly informed consent [86], but this requires providing the end-user with relevant information in an understandable form. This could be achieved by grouping the information across multiple layers [55, 145]. Our ranked list of privacy attributes serves as a basis for a user-centric privacy visualization that covers all important aspects of privacy.

5.3.3 PbD Guidelines Should Be More User-centric. One of the most striking findings was the fact that the two attributes rated as most important by both privacy experts and users (*sale* and *sharing*) were rarely covered by PbD principles. To avoid anxiety, uncertainty, or even fear [100], the gap between privacy concerns and guidelines aimed at addressing them must be reduced. PbD is aimed at taking the privacy concerns of the end-user into consideration during development, and so issues related to data sharing (in particular sale of user data) must be part of PbD guidelines. Ideally, since the lowest average importance rating was six on a 0-to-10 scale, PbD guidelines should cover all of the attributes on our list, with the possible exception of functionality. This is because functionality was ranked as one of the least important attributes and was sometimes marked as confusing by both privacy experts and users.

5.3.4 The Right to Be Forgotten Should Not Be Forgotten. The *right to be forgotten* was rarely mentioned in the PbD guidelines we reviewed. In 2014 however, the European Court of Justice ruled that European users can request the removal of personal data from online service providers and the GDPR mandates this as well (despite the fact that the right to be forgotten is not one of the GDPR's PbD principles). Newman questions whether the right to be forgotten is financially and legally feasible [98]. Still, according to Ausloos [17], the ability to demand the erasure of personal data can and must be available in data processing situations where consent was required and more widely assuming normative, economical, technical, and legislative changes. Even though most PbD guidelines already recommend obtaining consent (i.e., *control*) and recommend removal of data when it is no longer necessary (i.e., *retention*), the right to be forgotten goes a step further by giving users the ability to withdraw consent. Therefore, the right to be forgotten (or its diluted form, the "right to erasure" [11]), should be an integral part of future PbD guidelines.

5.4 Research Challenges

5.4.1 Structuring Privacy Policies. Privacy policies often focus on *collection*, *sale*, and *sharing* of user data, but our survey revealed that the *right to be forgotten* and *security* are of increasing concern. Furthermore, regulation increasingly mandates that privacy policies provide information about potential *disclosure* to (foreign) government entities, *accountability* in case of breaches, and the ability to *correct* one's data. The Unified List of Privacy Attributes of Section 4.3 is based on extensive review and comparison of privacy attributes covered by privacy visualizations and PbD guidelines aimed at online services in general. Therefore, it represents a complete and technology-/domain-independent checklist of aspects related to online privacy. A valuable research direction is to investigate whether such a checklist can be used to verify the completeness

of privacy policies [8], to structure (or even automatically restructure [161]) privacy policies, to make automatically generated privacy policies more readable [162], or to automatically analyze privacy policies [12].

5.4.2 Developing a Privacy Rating System. Similarly to PrivOnto [104], the privacy attributes on our list can be operationalized so that they can be used measure and compare the privacy level of online services on multiple metrics. The privacy attributes could also be used to (semi-)automatically annotate privacy policies [156]. In the long term, the privacy attributes could be used to produce or generate standardized, understandable, machine-readable summaries of privacy policies that enable both providers and users to assess, communicate, and compare the privacy of online services. To explore this direction, we started developing a free online service that implements some of these ideas: www.privacyrating.info. However, usability testing is critical, and the user testing performed on the DCI approach [147] and by Fox et al. [62] serves as a starting point for evaluating current and future proposals.

5.4.3 Investigating Context-dependency of Privacy Attributes. The Unified List of Privacy attributes in Section 4.3 is a first step toward a standardized list of privacy attributes that can function as the foundation of a privacy visualization. However, the work of Nissenbaum [100] and Martin [88] showed that information privacy is discriminate, embedded in the context, and based on a social contract between the various stakeholders involved in the information exchange. It seems that privacy perception is not universal, but depends on the contextual factors such as the type of data and the disclosure scenario [99, 113, 159]. For instance, the importance of some attributes might differ for an eHealth service compared to an online shopping website. Nevertheless, Solove [135] suggests a certain congruity between situations of personal data disclosure online.

We excluded mobile- and IoT-specific papers from our survey, because the domains are subject to different privacy concerns, and constrained in terms of privacy communication. Mobile apps are among the most privacy intrusive means of interacting with an online service [12] and privacy policies are almost always incorrect, incomplete, imprecise, inconsistent and/or privacy-unfriendly [160]. Even health apps are often not GDPR compliant [58].

The unified list of general attributes in Section 4.3 can be used as a reference and starting point for domain-, technology-, or target-group-specific guidelines or visualizations. However, the extent to which privacy is context dependent remains an open problem. Are specialized privacy labels needed or is a universal privacy visualization effective? How specific should PbD guidelines be?

6 CONCLUSIONS

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

Our study revealed that some attributes are considered important by both privacy experts and users: what type of personal data is collected, with whom it is shared with, and whether or not it is sold. The PbD guidelines we reviewed also emphasize collection, but mention purpose more often than sharing or sale. Furthermore, PbD guidelines often focus on ensuring information security and transparency while providing users with privacy controls. Privacy visualizations take a user-centric perspective, focusing on collection, purpose, and sharing. Overall, we see an increase in publications pertaining to PbD and privacy visualizations. The right to be forgotten and accountability of service providers are increasingly mentioned in both regulations and guidelines.

Both were found to be important in our survey. Disclosure to law enforcement, retention periods, and correctness of data are also mentioned increasingly often in publications covering online privacy, although these were ranked as relatively unimportant by our sample of privacy experts and users. Pseudonymization, anonymization, and the tradeoff between functionality and privacy are mentioned in a minority of the literature we reviewed and were perceived to be relatively unimportant by the users and privacy experts we surveyed.

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

REFERENCES

- [1] Technical Committee ISO/IEC JTC 1/SC 27. 2011. *ISO/IEC 29100:2011 Information Technology – Security Techniques – Privacy Framework*. Standard ISO/IEC 29100:2011(E). International Organization for Standardization and International Electrotechnical Commission, Geneva, Switzerland.
- [2] †Hezam A. Abdul-Ghani and Dimitri Konstantas. 2019. A comprehensive study of security and privacy guidelines, threats, and countermeasures: An IoT perspective. *J. Sens. Actuat. Netw.* 8, 2 (April 2019), 38 pages. <https://doi.org/10.3390/jsan8020022>
- [3] Mark S. Ackerman, Lorrie Faith Cranor, and Joseph Reagle. 1999. Privacy in e-commerce: Examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM Conference on Electronic Commerce (EC'99)*. ACM, New York, NY, 1–8. <https://doi.org/10.1145/336992.336995>
- [4] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie F. Cranor, Saranga Komanduri, Pedro G. Leon, Norman Sadeh, Florian Schaub, Many Sleeper, Yang Wang, and Shomir Wilson. 2017. Nudges for privacy and security: Understanding and assisting users' choices Online. *ACM Comput. Surv.* 50, 3, Article 44 (Oct. 2017), 41 pages. <https://doi.org/10.1145/3054926>
- [5] Alessandro Acquisti, Curtis Taylor, and Liad Wagman. 2016. The economics of privacy. *J. Econ. Lit.* 54, 2 (Jun. 2016), 442–492. <https://doi.org/10.1257/jel.54.2.442>
- [6] Amir S. Ahmadian, Daniel Strüder, and Jan Jürjens. 2019. Privacy-enhanced system design modeling based on privacy features. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC'19)*. ACM, New York, NY, 1492–1499. <https://doi.org/10.1145/3297280.3297431>
- [7] AICPA and CICA. 2009. *Generally Accepted Privacy Principles*. Reports 0001069. The American Institute of Certified Public Accountants, Inc. and The Canadian Institute of Chartered Accountants, Toronto, Canada.
- [8] Maryam Al-Jamal and Emad Abu-Shanab. 2015. Privacy policy of e-government websites: An itemized checklist proposed and tested. *Manag. Res. Pract.* 7, 3 (August 2015), 80–95. <https://doi.org/10.15849/icit.2015.0066>
- [9] Anita L. Allen. 2016. Protecting one's own privacy in a big data economy. *Harv. L. Rev. F.* 130 (December 2016), 71–78.
- [10] Majed Alshammari and Andrew Simpson. 2017. Towards a principled approach for engineering privacy by design. In *Privacy Technologies and Policy*, Erich Schweighofer, Herbert Leitold, Andreas Mittrakas, and Kai Rannenberg (Eds.). Lecture Notes in Computer Science (including subseries Security and Cryptology), Vol. 10518. Springer International Publishing AG, Cham, Switzerland, 161–177. https://doi.org/10.1007/978-3-319-67280-9_9
- [11] Meg L. Ambrose and Jef Ausloos. 2013. The right to be forgotten across the pond. *J. Inf. Policy* 3 (2013), 1–23. <https://doi.org/10.5325/jinfopoli.3.2013.0001>
- [12] Benjamin Andow, Samin Yaseer Mahmud, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Serge Egelman. 2020. Actions speak louder than words: Entity-sensitive privacy policy and data flow analysis with polichex. In *Proceedings of the 29th Security Symposium (USENIX'20)*. USENIX Association, Berkeley, CA, 985–1002.
- [13] Sheila F. Anthony. 2001. The Case for Standardization of Privacy Policy Formats. Retrieved June, 1, 2021 from <https://www.ftc.gov/public-statements/2001/07/case-standardization-privacy-policy-formats>.
- [14] Annie I. Antón, Elisa Bertino, Ninghui Li, and Ting Yu. 2007. A roadmap for comprehensive online privacy policy management. *Commun. ACM* 50, 7 (July 2007), 109–116. <https://doi.org/10.1145/1272516.1272522>
- [15] Annie I. Antón, Julia Brande Earp, Qingfeng He, William Stufflebeam, Davide Bolchini, and Carlos Jensen. 2004. Financial privacy policies and the need for standardization. *IEEE Secur. Priv.* 2, 2 (August 2004), 36–45. <https://doi.org/10.1109/MSECP.2004.1281243>
- [16] Memoona J. Anwar, Asif Q. Gill, and Ghassan Beydoun. 2018. A review of information privacy laws and standards for secure digital ecosystems. In *Proceedings of the 29th Australasian Conference on Information Systems (ACIS'18)*. Australasian Association for Information Systems, Sydney, Australia. <https://doi.org/10.5130/acis2018.bb>

- [17] Jef Ausloos. 2012. The ‘right to be forgotten’—worth remembering? *Comput. Law Secur. Rev.* 28, 2 (April 2012), 143–152. <https://doi.org/10.1016/j.clsr.2012.01.006>
- [18] Gökhan Bal. 2014. Designing privacy indicators for smartphone app markets: A new perspective on the nature of privacy risks of apps. In *Proceedings of the 20th Americas Conference on Information Systems (AMCIS’14)*. Association for Information Systems.
- [19] Maria T. Baldassarre, Vita S. Barletta, Danilo Caivano, and Michele Scalera. 2019. Privacy oriented software development. In *Quality of Information and Communications Technology*, Mario Piattini, Paulo Rupino da Cunha, Ignacio García Rodríguez de Guzmán, and Ricardo Pérez-Castillo (Eds.). Communications in Computer and Information Science, Vol. 1010. Springer Nature Switzerland AG, Cham, Switzerland, 18–32. https://doi.org/10.1007/978-3-030-29238-6_2
- [20] Gaurav Bansal, Fatemeh “Mariam” Zahedi, and David Gefen. 2010. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decis. Support Syst.* 49, 2 (May 2010), 138–150. <https://doi.org/10.1016/j.dss.2010.01.010>
- [21] †Masoud Barati and Omer Rana. 2020. Enhancing user privacy in IoT: Integration of GDPR and blockchain. In *Blockchain and Trustworthy Systems*, Zibin Zheng, Hong-Ning Dai, Mingdong Tang, and Xiangping Chen (Eds.). Communications in Computer and Information Science, Vol. 1156. Springer Nature Singapore Pte Ltd., Singapore, 322–335. https://doi.org/10.1007/978-981-15-2777-7_26
- [22] Luca Belli, Molly Schwartz, and Luiza Louzada. 2017. Selling your soul while negotiating the conditions: From notice and consent to data control by design. *Health Technol.* 7, 4 (Mar. 2017), 453–467. <https://doi.org/10.1007/s12553-017-0185-3>
- [23] Steven Bellman, Eric J. Johnson, Stephen J. Kobrin, and Gerald L. Lohse. 2004. International differences in information privacy concerns: A global survey of consumers. *Inf. Soc.* 20, 5 (Aug. 2004), 313–324. <https://doi.org/10.1080/01972240490507956>
- [24] Kevin Benton, L. Jean Camp, and Vaibhav Garg. 2013. Studying the effectiveness of android application permissions requests. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops’13)*. IEEE, New York, NY, 291–296. <https://doi.org/10.1109/PerComW.2013.6529497>
- [25] Konstantin Beznosov, Philip Inglesant, Jorge Lobo, Rob Reeder, and Mary E. Zurko. 2009. Usability meets access control: Challenges and research opportunities. In *Proceedings of the 14th ACM Symposium on Access Control Models and Technologies (SACMAT’09)*. ACM, New York, NY, 73–74. <https://doi.org/10.1145/1542207.1542220>
- [26] Christoph Bier, Pascal Birnstill, Erik Krempel, Hauke Vagts, and Jürgen Beyerer. 2014. Enhancing privacy by design from a developer’s perspective. In *Privacy Technologies and Policy*, Bart Preneel and Demosthenes Ikononou (Eds.). Lecture Notes in Computer Science, Vol. 8319. Springer-Verlag, Berlin, 73–85. https://doi.org/10.1007/978-3-642-54069-1_5
- [27] †Giorgia Bincoletto. 2019. A data protection by design model for privacy management in electronic health records. In *Privacy Technologies and Policy*, Maurizio Naldi, Giuseppe F. Italiano, Kai Rannenberg, Manel Medina, and Athena Bourka (Eds.). Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 11498. Springer Nature Switzerland AG, Cham, Switzerland, 161–181. https://doi.org/10.1007/978-3-030-21752-5_11
- [28] †Ian Brown. 2014. Britain’s smart meter programme: A case study in privacy by design. *Int. Rev. Law Comput. Technol.* 28, 2 (Jul. 2014), 172–184. <https://doi.org/10.1080/13600869.2013.801580>
- [29] The Privacy Company B.V. 2020. Privacy by Design Framework. (January 2020). Retrieved August 10, 2020 from https://uploads-ssl.webflow.com/5d5d0a009052fec16249aaab/5dc3e8a351595bfff2a94fda_Privacy%20by%20Design%20framework%20V3.pdf.
- [30] Rose L. Casey. 2016. *Privacy by Design: Setting a New Standard for Privacy Certification*. Reports 15-2971-H. Deloitte LLP, Ontario, Canada.
- [31] Fred H. Cate. 1997. *Privacy in the Information Age*. Brookings Institution Press, Washington, DC.
- [32] Ann Cavoukian. 2006. *Creation of a Global Privacy Standard*. Report. Information and Privacy Commissioner of Ontario, Ontario, Canada.
- [33] Ann Cavoukian. 2009. *Privacy by Design: The 7 Foundational Principles*. Report. Information and Privacy Commissioner of Ontario, Ontario, Canada.
- [34] Ann Cavoukian. 2010. *Privacy by Design: The 7 Foundational Principles: Implementation and Mapping of Fair Information Practices*. Report. Information and Privacy Commissioner of Ontario, Ontario, Canada.
- [35] †Ann Cavoukian. 2011. Privacy by design: Best practices for privacy and the smart grid. In *ISSE 2010 Securing Electronic Business Processes*, Norbert Pohlmann, Helmut Reimer, and Wolfgang Schneider (Eds.). Vieweg+Teubner, Wiesbaden, Germany, 260–270. https://doi.org/10.1007/978-3-8348-9788-6_25
- [36] Ann Cavoukian. 2012. Privacy by design [leading edge]. *IEEE Technol. Soc. Mag.* 31, 4 (December 2012), 18–19. <https://doi.org/10.1109/MTS.2012.2225459>

- [37] Ann Cavoukian. 2020. Understanding how to implement privacy by design, one step at a time. *IEEE Consum. Electron. Mag.* 9, 2 (March 2020), 78–82. <https://doi.org/10.1109/MCE.2019.2953739>
- [38] †Ann Cavoukian and Michelle Chibba. 2016. Cognitive cities, big data and citizen participation: The essentials of privacy and security. In *Towards Cognitive Cities*, Edy Portmann and Matthias Finger (Eds.). Studies in Systems, Decision and Control, Vol. 63. Springer International Publishing AG, Cham, Switzerland, 61–82. https://doi.org/10.1007/978-3-319-33798-2_4
- [39] Yi Chen, Mingming Zha, Nan Zhang, Dandan Xu, Qianqian Zhao, Xuan Feng, Kan Yuan, Fnu Suya, Yuan Tian, Kai Chen, XiaoFeng Wang, and Wei Zou. 2019. Demystifying hidden privacy settings in mobile apps. In *Proceedings of the IEEE Symposium on Security and Privacy (SP'19)*. IEEE Computer Society, Los Alamitos, CA, 850–866. <https://doi.org/10.1109/SP.2019.00054>
- [40] Hanbyul Choi, Jonghwa Park, and Yoonhyuk Jung. 2018. The role of privacy fatigue in online privacy behavior. *Comput. Hum. Behav.* 81 (April 2018), 42–51. <https://doi.org/10.1016/j.chb.2017.12.001>
- [41] Michael Colesky and Julio C. Caiza. 2018. A system of privacy patterns for informing users: Creating a pattern system. In *Proceedings of the 23rd European Conference on Pattern Languages of Programs (EuroPLOP'18)*. ACM, New York, NY, Article 16, 11 pages. <https://doi.org/10.1145/3282308.3282325>
- [42] Michael Colesky, Julio C. Caiza, José M. Del Álamo, Jaap-Henk Hoepman, and Yod-Samuel Martin. 2018. A system of privacy patterns for user control. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC'18)*. ACM, New York, NY, 1150–1156. <https://doi.org/10.1145/3167132.3167257>
- [43] Michael Colesky, Jaap-Henk Hoepman, and Christiaan Hillen. 2016. A critical analysis of privacy design strategies. In *Proceedings of the 2016 IEEE Symposium on Security and Privacy Workshops (SPW'16)*. IEEE, New York, NY, 33–40. <https://doi.org/10.1109/SPW.2016.23>
- [44] Liane Colonna. 2013. Prism and the european union's data protection directive. *J. Marsh. J. Info. Tech. Priv. L.* 30, 2, Article 1 (2013), 27 pages.
- [45] Asia-Pacific Economic Cooperation. 2005. *APEC Privacy Framework*. Reports APEC#205-SO-01.2. Asia Pacific Economic Cooperation Secretariat, Singapore.
- [46] Lorrie Cranor. 2009. Find Web Sites That Respect Your Privacy. Retrieved December 10, 2019 from <http://www.privacybird.org>.
- [47] Digital Advertising Alliance. 2016. Put the Your AdChoices Icon to Work for You. Retrieved December 10, 2019 from <http://youradchoices.com/learn>.
- [48] †Nigel Davies and Marc Langheinrich. 2013. Privacy by design [From the editor in chief]. *IEEE Pervas. Comput.* 12, 2 (April 2013), 2–4. <https://doi.org/10.1109/MPRV.2013.34>
- [49] *André de Lima Salgado, Felipe Silva Dias, João Pedro Rodrigues Mattos, Renata Pontin de Mattos Fortes, and Patrick C. K. Hung. 2019. Smart toys and children's privacy: Usable privacy policy insights from a card sorting experiment. In *Proceedings of the 37th ACM International Conference on the Design of Communication (SIGDOC'19)*. ACM, New York, NY, Article 16, 8 pages. <https://doi.org/10.1145/3328020.3353951>
- [50] Michelle Finneran Denney, Jonathan Fox, and Thomas R. Finneran. 2014. *The Privacy Engineer's Manifesto: Getting from Policy to Code to QA to Value*. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4302-6356-2>
- [51] Michele Drgon, Gail Magnuson, and John Sabo. 2016. *Privacy Management Reference Model and Methodology (PMRM) Version 1.0*. Standards Track Work Product PMRM-v1.0-cs02. OASIS Committee Specification 02. OASIS Open 2016, Burlington, MA.
- [52] Olha Drozd. 2016. Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process. In *Privacy and Identity Management. Time for a Revolution?* David Aspinall, Jan Camenisch, Marit Hansen, Simone Fischer-Hübner, and Charles Raab (Eds.). IFIP Advances in Information and Communication Technology, Vol. 476. Springer International Publishing AG, Cham, Switzerland, 129–140. https://doi.org/10.1007/978-3-319-41763-9_9
- [53] Benjamin Edelman. 2009. Adverse selection in online “trust” certifications. In *Proceedings of the 11th International Conference on Electronic Commerce (ICEC'09)*. ACM, New York, NY, 205–212. <https://doi.org/10.1145/1593254.1593286>
- [54] Benjamin Edwards, Steven Hofmeyr, and Stephanie Forrest. 2016. Hype and heavy tails: A closer look at data breaches. *J. Cybersecur.* 2, 1 (December 2016), 3–14. <https://doi.org/10.1093/cybsec/tyw003>
- [55] Lilian Edwards and Wiebke Abel. 2014. *The Use of Privacy Icons and Standard Contract Terms for Generating Consumer Trust and Confidence in Digital Services*. CREATE Working Paper Series 2014/15 (October 2014). University of Glasgow, Glasgow, UK. <https://doi.org/10.5281/zenodo.12506>
- [56] *Serge Egelman, Raghudeep Kannavara, and Richard Chow. 2015. Is this thing on? Crowdsourcing privacy indicators for ubiquitous sensing platforms. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15)*. ACM, New York, NY, 1669–1678. <https://doi.org/10.1145/2702123.2702251>
- [57] *Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. 2019. Exploring how privacy and security factor into IoT device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI'19)*. ACM, New York, NY, Article 534, 12 pages. <https://doi.org/10.1145/3290605.3300764>

- [58] Ming Fan, Le Yu, Sen Chen, Hao Zhou, Xiapu Luo, Shuyue Li, Yang Liu, Jun Liu, and Ting Liu. 2020. An empirical evaluation of GDPR compliance violations in Android mHealth apps. In *IEEE 31st International Symposium on Software Reliability Engineering (ISSRE)*. IEEE, New York, NY, 253–264. <https://doi.org/10.1109/ISSRE5003.2020.00032>
- [59] Jan Fernback and Zizi Papacharissi. 2007. Online privacy as legal safeguard: The relationship among consumer, online portal, and privacy policies. *New Media Soc.* 9, 5 (October 2007), 715–734. <https://doi.org/10.1177/1461444807080336>
- [60] Simone Fischer-Hübner and Harald Zwingelberg. 2010. UI Prototypes: Policy Administration and Presentation Version 2. PrimeLife Project Deliverable D.4.3.2. (2010). Retrieved June 1, 2021 from <http://primelife.ercim.eu/>.
- [61] †Department for Digital Culture Media & Sport. 2018. *Secure by Design: Improving the Cyber Security of Consumer Internet of Things Report*. Policy Paper. UK Government, London, UK.
- [62] Grace Fox, Colin Tonge, Theo Lynn, and John Mooney. 2018. Communicating compliance: Developing a GDPR privacy label. In *Proceedings of the 24th Americas Conference on Information Systems (AMCIS'18)*. Association for Information Systems.
- [63] Gert Franke, Thomas Clever, Wouter van Dijk, Jeremy Raider, and Roel de Jonge. 2019. Privacy Label. Blog series. (November 2019). Retrieved December 10, 2019 from <https://medium.com/sensor-lab/the-privacy-illusion-994ed98ec3ab>.
- [64] Robert Gellman. 2019. Fair information practices: A basic history - Version 2.19. SSRN, (Oct. 2019). <https://doi.org/10.2139/ssrn.2415020>
- [65] *Martin Gisch, Alexander De Luca, and Markus Blanchebarbe. 2007. The privacy badge: A privacy-awareness user interface for small devices. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology (Mobility'07)*. ACM, New York, NY, 583–586. <https://doi.org/10.1145/1378063.1378159>
- [66] Joshua Gomez, Travis Pinnick, and Ashkan Soltani. 2009. Privacy Coding Methodology. Retrieved December 10, 2019 from http://knowprivacy.org/policies_methodology.html.
- [67] Canadian Standards Association (CSA Group). 2014. Model Code for the Protection of Personal Information of 23 December 2014, First Published in March 1996; Reaffirmed 2001. CAN/CSA-Q830-96. (December 2014). Retrieved December 10, 2019 from https://www.afn.ca/uploads/files/nihbforum/info_and_privacy_doc_-_csa_model_code_for_the_protection_of_personal_information.pdf.
- [68] Kevin Haninger and Kimberly M. Thompson. 2004. Content and ratings of teen-rated video games. *J. Am. Med. Assoc.* 291, 7 (February 2004), 856–865. <https://doi.org/10.1001/jama.291.7.856>
- [69] Marit Hansen. 2009. Putting privacy pictograms into practice—A European perspective. In *Im Focus das Leben, Beiträge der 39. Jahrestagung der Gesellschaft für Informatik e.V. (INFORMATIK'09)*, Stefan Fischer, Erik Maehle, and Rüdiger Reischuk (Eds.). Lecture Notes in Informatics, Vol. P-154. Gesellschaft für Informatik, Bonn, Germany, 1703–1716.
- [70] Marit Hansen, Ari Schwartz, and Alissa Cooper. 2008. Privacy and identity management. *IEEE Secur. Priv.* 6, 2 (April 2008), 38–45. <https://doi.org/10.1109/MSP.2008.41>
- [71] Kirstie Hawkey and Kori M. Inkpen. 2007. PrivateBits: Managing visual privacy in web browsers. In *Proceedings of the Graphics Interface Conference (GI'07)*. ACM, New York, NY, 215–223. <https://doi.org/10.1145/1268517.1268553>
- [72] Jaap-Henk Hoepman. 2014. Privacy design strategies (Extended abstract). In *ICT Systems Security and Privacy Protection*, Nora Cuppens-Boulahia, Frédéric Cuppens, Sushil Jajodia, Anas Abou El Kalam, and Thierry Sans (Eds.). IFIP Advances in Information and Communication Technology, Vol. 428. Springer-Verlag, Berlin, 446–459. https://doi.org/10.1007/978-3-642-55415-5_38
- [73] Christian P. Hoffmann, Christoph Lutz, and Giulia Ranzini. 2016. Privacy cynicism: A new approach to the privacy paradox. *Cyberpsychology* 10, 4, Article 7 (December 2016), 18 pages. <https://doi.org/10.5817/CP2016-4-7>
- [74] Leif-Erik Holtz, Katharina Nocun, and Marit Hansen. 2011. Towards displaying privacy information with icons. In *Privacy and Identity Management for Life*, Simone Fischer-Hübner, Penny Duquenoy, Marit Hansen, Ronald Leenes, and Ge Zhang (Eds.). IFIP Advances in Information and Communication Technology, Vol. 352. Springer-Verlag, Berlin, 338–348. https://doi.org/10.1007/978-3-642-20769-3_27
- [75] *Renato Iannella and Adam Finden. 2009. Privacy awareness: Icons and expression for social networks. In *Proceedings of the 8th International Workshop for Technical, Economic and Legal Aspects of Business Models for Virtual Goods, Incorporating the 6th International Open Digital Rights Language Workshop (ODRL'09)*. Presses universitaires de Namur, Namur, Belgium, 13 pages.
- [76] *Shane D. Johnson, John M. Blythe, Matthew Manning, and Gabriel T. W. Wong. 2020. The impact of IoT security labelling on consumer product choice and willingness to pay. *PLoS One* 15, 1 (January 2020), 1–21. <https://doi.org/10.1371/journal.pone.0227800>
- [77] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. 2009. A “nutrition label” for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS'09)*. ACM, New York, NY, Article 4, 12 pages. <https://doi.org/10.1145/1572532.1572538>

- [78] Patrick Gage Kelley, Sunny Consolvo, Lorrie F. Cranor, Jaeyeon Jung, Norman Sadeh, and David Wetherall. 2012. A conundrum of permissions: Installing applications on an Android smartphone. In *Financial Cryptography and Data Security*, Jim Blyth, Sven Dietrich, and L. Jean Camp (Eds.). Lecture Notes in Computer Science, Vol. 7938. Springer-Verlag, Berlin, 68–79. https://doi.org/10.1007/978-3-642-34638-5_6
- [79] Anne Klinefelter. 2011. When to research is to reveal: The growing threat to attorney and client confidentiality from online tracking. *Va. J. L. Tech.* 16, 1 (2011), 1.
- [80] Linda Kool, Jelte Timmer, Lambèr M. M. Royakkers, and Q.C. (Rinie) van Est. 2017. *Urgent Upgrade: Protect Public Values in Our Digitized Society*. Report. Rathenau Instituut, The Hague, Netherlands.
- [81] Anastasia Kozyreva, Stefan Herzog, Philipp Lorenz-Spreen, Ralph Hertwig, and Stephan Lewandowsky. 2020. *Artificial Intelligence in Online Environments: Representative Survey of Public Attitudes in Germany*. Survey Report. Max Planck Institute for Human Development, Munich, Germany. <https://doi.org/10.17617/2.3188061>
- [82] †Marc Langheinrich. 2001. Privacy by design - Principles of privacy-aware ubiquitous systems. In *Proceedings of the Annual Conference on Ubiquitous Computing (Ubicomp'01)*, Gregory D. Abowd, Barry Brumitt, and Steven Shafer (Eds.). Lecture Notes in Computer Science, Vol. 2201. Springer-Verlag, Berlin, 273–291. https://doi.org/10.1007/3-540-45427-6_23
- [83] Stephen E. Levy and Carl Gutwin. 2005. Improving understanding of website privacy policies with fine-grained policy anchors. In *Proceedings of the 14th International Conference on World Wide Web (WWW'05)*. ACM, New York, NY, 480–488. <https://doi.org/10.1145/1060745.1060816>
- [84] Yuanchun Li, Fanglin Chen, Toby Jia-Jun Li, Yao Guo, Gang Huang, Matthew Fredrikson, Yuvraj Agarwal, and Jason I. Hong. 2017. PrivacyStreams: Enabling transparency in personal data processing for mobile apps. *Proc. ACM Interact. Mob. Wear. Ubiquit. Technol.* 1, 3, Article 76 (September 2017), 26 pages. <https://doi.org/10.1145/3130941>
- [85] Eleni-Laskarina Makri and Costas Lambrinouidakis. 2015. Privacy principles: Towards a common privacy audit methodology. In *Trust, Privacy and Security in Digital Business*, Simone Fischer-Hübner, Costas Lambrinouidakis, and Javier López (Eds.). Lecture Notes in Computer Science (including subseries Security and Cryptology), Vol. 9264. Springer International Publishing AG, Cham, Switzerland, 219–234. https://doi.org/10.1007/978-3-319-22906-5_17
- [86] James Martin and Nicolas Christin. 2016. Ethics in cryptomarket research. *Int. J. Drug Policy* 35 (September 2016), 84–91. <https://doi.org/10.1016/j.drugpo.2016.05.006>
- [87] Kirsten Martin. 2016. Data aggregators, consumer data, and responsibility online: Who is tracking consumers online and should they stop? *Inf. Soc.* 32, 1 (December 2016), 51–63. <https://doi.org/10.1080/01972243.2015.1107166>
- [88] Kirsten Martin. 2016. Understanding privacy online - development of a social contract approach to privacy. *J. Bus. Ethics* 137, 3 (February 2016), 551–569. <https://doi.org/10.1007/s10551-015-2565-9>
- [89] Kirsten Martin and Katie Shilton. 2016. Putting mobile application privacy in context: An empirical study of user privacy expectations for mobile devices. *Inf. Soc.* 32, 3 (April 2016), 200–216. <https://doi.org/10.1080/01972243.2016.1153012>
- [90] Yod-Samuel Martin, Jose M. del Alamo, and Juan C. Yelmo. 2014. Engineering privacy requirements valuable lessons from another realm. In *Proceedings of the IEEE 1st International Workshop on Evolving Security and Privacy Requirements Engineering (ESPRE'14)*. IEEE, New York, NY, 19–24. <https://doi.org/10.1109/ESPRE.2014.6890523>
- [91] Yod-Samuel Martin and José M. del Álamo. 2017. A metamodel for privacy engineering methods. In *Proceedings of the 3rd International Workshop on Privacy Engineering co-located with 38th IEEE Symposium on Security and Privacy (S&P'17)*, José M. del Álamo, Seda F. Gürses, and Anupam Datta (Eds.). CEUR Workshop Proceedings, Vol. 1873. CEUR-WS Team, Aachen, Germany, 41–48.
- [92] Kate Mayfield. 2016. Pseudonymisation: A 20-yearold idea never seemed so timely. *J. Direct Data Digit. Mark. Pract.* 17 (September 2016), 222–226. <https://doi.org/10.1057/s41263-016-0005-x>
- [93] Miriam J. Metzger. 2007. Communication privacy management in electronic commerce. *J. Comput.-Mediat. Commun.* 12, 2 (January 2007), 335–361. <https://doi.org/10.1111/j.1083-6101.2007.00328.x>
- [94] Arthur R. Miller. 1969. Personal privacy in the computer age: The challenge of a new technology in an information-oriented society. *Mich. Law. Rev.* 67, 6 (April 1969), 1089–1246. <https://doi.org/10.2307/1287516>
- [95] Miguel E. Morales-Trujillo, Erick O. Matla-Cruz, Gabriel A. García-Mireles, and Mario Piattini. 2018. Privacy by design in software engineering: A systematic mapping study. In *Proceedings of the 21st Iberoamerican Conference on Software Engineering (CIBSE'18)*, Rubby Casallas, Kelly Garces, and Mario Sanchez (Eds.). Curran Associates, Inc., Red Hook, NY, 107–120. <https://doi.org/10.19153/cleiej.22.1.4>
- [96] Victor Morel and Raúl Pardo. 2020. SoK: Three facets of privacy policies. In *Proceedings of the 19th Workshop on Privacy in the Electronic Society (WPES'20)*. ACM, New York, NY, 41–56. <https://doi.org/10.1145/3411497.3420216>
- [97] Vivian Genaro Motti and Kelly Caine. 2016. Towards a visual vocabulary for privacy concepts. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 60, 1 (Sep. 2016), 1078–1082. <https://doi.org/10.1177/1541931213601249>
- [98] Abraham L. Newman. 2015. What the “right to be forgotten” means for privacy in a digital age. *Science* 347, 6221 (January 2015), 507–508. <https://doi.org/10.1126/science.aaa4603>

- [99] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79, 1 (January 2004), 119–158.
- [100] Helen Nissenbaum. 2011. A contextual approach to privacy online. *Daedalus* 140, 4 (September 2011), 32–48. https://doi.org/10.1162/DAED_a_00113
- [101] OECD. 2013. *The OECD Privacy Framework*. Booklet. Organisation for Economic Co-operation and Development, Paris, France.
- [102] Office of the Australian Information Commissioner (OAIC). 2014. *Australian Privacy Principles*. A Summary for APP entities from 12 March 2014. Australian Government, Sydney, Australia. Retrieved December 10, 2019 from <https://www.oaic.gov.au/assets/privacy/guidance-and-advice/app-quick-reference-tool.pdf>.
- [103] Office of the Privacy Commissioner of Canada (OAIC). 2000. *Personal Information Protection and Electronic Documents Act (PIPEDA)*. Act S.C. 200, c.5. Current to July 28, 2020, last amended on June 21, 2019. Minister of Justice, Ontario, Canada. Retrieved July 17, 2020 from <https://laws-lois.justice.gc.ca/eng/acts/P-8.6/>.
- [104] Alessandro Oltramari, Dhivya Piraviperumal, Florian Schaub, Shomir Wilson, Sushain Chervirala, Thomas B. Norton, N. Cameron Russell, Peter Story, Joel Reidenberg, and Norman Sadeh. 2018. PrivOnto: A semantic framework for the analysis of privacy policies. *Semantic Web J.* 9, 2 (January 2018), 185–203. <https://doi.org/10.3233/SW-170283>
- [105] Luci Pangrazio and Neil Selwyn. 2019. ‘Personal data literacies’: A critical literacies approach to enhancing understandings of personal digital data. *New Media Soc.* 21, 2 (September 2019), 419–437. <https://doi.org/10.1177/1461444818799523>
- [106] The European Parliament and the Council of European Union. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)*. Legislation OJ L 119, 2.5. Publications Office of the European Union, Luxembourg, Luxembourg.
- [107] Paul A. Pavlou. 2011. State of the information privacy literature: Where are we now and where should we go? *Manage. Inf. Syst. Quart.* 35, 4 (December 2011), 977–988. <https://doi.org/10.2307/41409969>
- [108] Pedram Pedarsani and Matthias Grossglauser. 2011. On the privacy of anonymized networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD’11)*. ACM, New York, NY, 1235–1243. <https://doi.org/10.1145/2020408.2020596>
- [109] †Juanita Pedraza, Miguel A. Patricio, Agustín De Asís, and José M. Molina. 2011. Regulatory model for AAL. In *Soft Computing Models in Industrial and Environmental Applications, 6th International Conference (SOCO’11)*, Emilio Corchado, Václav Snášel, Javier Sedano, Aboul E. Hassanien, José L. Calvo, and Dominik Ślezak (Eds.). Advances in Intelligent and Soft Computing, Vol. 87. Springer-Verlag, Berlin, 183–192. https://doi.org/10.1007/978-3-642-19644-7_20
- [110] †Juanita Pedraza, Miguel A. Patricio, Agustín de Asís, and José M. Molina. 2013. Privacy-by-design rules in face recognition system. *Neurocomputing* 109, 3 (Jun. 2013), 49–55. <https://doi.org/10.1016/j.neucom.2012.03.023>
- [111] †Charith Perera, Mahmoud Barhamgi, Arosha K. Bandara, Muhammad Ajmal, Blaine Price, and Bashar Nuseibeh. 2020. Designing privacy-aware internet of things applications. *Inf. Sci.* 512 (Feb. 2020), 238–257. <https://doi.org/10.1016/j.ins.2019.09.061>
- [112] †Charith Perera, Ciaran McCormick, Arosha K. Bandara, Blaine A. Price, and Bashar Nuseibeh. 2016. Privacy-by-Design framework for assessing Internet of Things applications and platforms. In *Proceedings of the 6th International Conference on the Internet of Things (IoT’16)*. ACM, New York, NY, 83–92. <https://doi.org/10.1145/2991561.2991566>
- [113] Joseph Phelps, Glen Nowak, and Elizabeth Ferrell. 2000. Privacy concerns and consumer willingness to provide personal information. *J. Publ. Policy Mark.* 19, 1 (April 2000), 27–41. <https://doi.org/10.1509/jppm.19.1.27.16941>
- [114] †Denis Pinkas. 2016. An eID mechanism built along privacy by design principles using secure elements, pseudonyms and attributes. In *Open Identity Summit 2016 der Gesellschaft für Informatik e.V. (GI’16)*, Detlef Hühnlein, Heiko Roßnagel, Christian H. Schunck, and Maurizio Talamo (Eds.). LNI, Vol. P-264. Gesellschaft für Informatik e.V., Bonn, 93–104.
- [115] Travis Pinnick. 2011. Privacy Short Notice Design. Retrieved December 10, 2019 from <https://www.trustarc.com/blog/?p=1253>.
- [116] *Kellie Poneres, Foad Hamidi, Aaron Massey, and Amy Hurst. 2018. Using icons to communicate privacy characteristics of adaptive assistive technologies. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS’18)*. ACM, New York, NY, 388–390. <https://doi.org/10.1145/3234695.3241003>
- [117] †Davy Preuveneers, Wouter Joosen, and Elisabeth Ilie-Zudor. 2016. Data protection compliance regulations and implications for smart factories of the future. In *Proceedings of the 12th International Conference on Intelligent Environments (IE’16)*. IEEE Computer Society, Los Alamitos, CA, 40–47. <https://doi.org/10.1109/IE.2016.15>
- [118] Hannah Quay-de la Vallee, Paige Selby, and Shriram Krishnamurthi. 2016. On a (per)mission: Building privacy into the app marketplace. In *Proceedings of the 6th Workshop on Security and Privacy in Smartphones and Mobile (SPSM’16)*. ACM, New York, NY, 63–72. <https://doi.org/10.1145/2994459.2994466>

- [119] Kopo M. Ramokapane, Anthony C. Mazeli, and Awais Rashid. 2019. Skip, skip, skip, accept!!!: A study on the usability of smartphone manufacturer provided default features and user privacy. *Proc. Priv. Enhanc. Technol.* 2019, 2 (May 2019), 209–227. <https://doi.org/10.2478/popets-2019-0027>
- [120] Sandra D. Ringmann, Hanno Langweg, and Marcel Waldvogel. 2018. Requirements for legally compliant software based on the GDPR. In *On the Move to Meaningful Internet Systems*, Hervé Panetto, Christophe Debruyne, Henderik A. Proper, Claudio Agostino Ardagna, Dumitru Roman, and Robert Meersman (Eds.). Lecture Notes in Computer Science (including subseries Programming and Software Engineering), Vol. 11230. Springer Nature Switzerland AG, Cham, Switzerland, 258–276. https://doi.org/10.1007/978-3-030-02671-4_15
- [121] Anna Romanou. 2018. The necessity of the implementation of privacy by design in sectors where data protection concerns arise. *Comput. Law Secur. Rev.* 34, 1 (Feb. 2018), 99–110. <https://doi.org/10.1016/j.clsr.2017.05.021>
- [122] Arianna Rossi and Monica Palmirani. 2017. A visualization approach for adaptive consent in the European data protection framework. In *Proceedings of the 7th International Conference for E-Democracy and Open Government (CeDEM'17)*, Peter Parycek and Noella Edelmann (Eds.). IEEE Computer Society, Los Alamitos, CA, 159–170. <https://doi.org/10.1109/CeDEM.2017.23>
- [123] Arianna Rossi and Monica Palmirani. 2019. DAPIS: An ontology-based data protection icon set. In *Knowledge of the Law in the Big Data Age*, Ginevra Peruginelli and Sebastiano Faro (Eds.). IOS Press BV, Amsterdam, Netherlands, 181–195. <https://doi.org/10.3233/FAIA317>
- [124] †Guilda Rostama, Alexandre Bekhradi, and Bernard Yannou. 2017. From privacy by design to design for privacy. In *Proceedings of the 21st International Conference on Engineering Design*, Anja Maier, Stanko Škec, Harrison Kim, Michael Kokkolaras, Josef Oehmen, Georges Fadel, Filippo Salustri, and Mike Van der Loos (Eds.). ICED17, Vol. DS87–6. Design Society, Glasgow, United Kingdom, 317–326.
- [125] Jeffrey Rothfeder. 1992. *Privacy for Sale: How Computerization Has Made Everyone's Private Life an Open Secret*. Simon & Schuster Trade, New York, NY.
- [126] John T. Sabo. 2007. ISTPA operational analysis of international privacy requirements. In *ISSE/SECURE 2007 Securing Electronic Business Processes*, Norbert Pohlmann, Helmut Reimer, and Wolfgang Schneider (Eds.). Friedr. Vieweg & Sohn Verlag, Wiesbaden, Germany, 18–25. https://doi.org/10.1007/978-3-8348-9418-2_2
- [127] Peter Schaar. 2010. Privacy by design. *Ident. Inf. Soc.* 3, 2 (April 2010), 267–274. <https://doi.org/10.1007/s12394-010-0055-x>
- [128] Gerardo Schneider. 2018. Is privacy by construction possible? In *Leveraging Applications of Formal Methods, Verification and Validation. Modeling*, Tiziana Margaria and Bernhard Steffen (Eds.). Lecture Notes in Computer Science (including subseries Theoretical Computer Science and General Issues), Vol. 11244. Springer Nature Switzerland AG, Cham, Switzerland, 471–485. https://doi.org/10.1007/978-3-030-03418-4_28
- [129] Bruce Schneier. 2015. *Data and Goliath: The Hidden Battles to Capture Your Data and Control Your World* (1st. ed.). W. W. Norton & Company, New York, NY.
- [130] †Elaine Sedenberg, John Chuang, and Deirdre Mulligan. 2016. Designing commercial therapeutic robots for privacy preserving systems and ethical research practices within the home. *Int. J. Soc. Robot.* 8 (Jun. 2016), 575–587. <https://doi.org/10.1007/s12369-016-0362-y>
- [131] *Yun Shen and Pierre-Antoine Vervier. 2019. IoT security and privacy labels. In *Privacy Technologies and Policy*, Maurizio Naldi, Giuseppe F. Italiano, Kai Rannenberg, Manel Medina, and Athena Bourka (Eds.). Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 11498. Springer Nature Switzerland AG, Cham, Switzerland, 136–147. https://doi.org/10.1007/978-3-030-21752-5_9
- [132] Andy P. Siddaway, Alex M. Wood, and Larry V. Hedges. 2019. How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annu. Rev. Psychol.* 70 (Jan. 2019), 747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- [133] H. Jeff Smith, Tamara Dinev, and Heng Xu. 2011. Information privacy research: An interdisciplinary review. *Manage. Inf. Syst. Q.* 35, 4 (December 2011), 989–1015. <https://doi.org/10.2307/41409970>
- [134] *Karen L. Smith, Leslie R. Shade, and Tamara Shepherd. 2017. Open privacy badges for digital policy literacy. *Int. J. Commun.* 11 (Nov. 2017), 2784–2805.
- [135] Daniel J. Solove. 2002. Conceptualizing privacy. *Calif. L. Rev.* 90, 4 (2002), 1087–1155.
- [136] Daniel J. Solove. 2008. *Understanding Privacy*. Legal Studies Research Paper No. 420, GWU Law School Public Law Research Paper No. 420. Harvard University Press, Cambridge, MA.
- [137] Daniel J. Solove, Marc Rotenberg, and Paul M. Schwartz. 2006. *Privacy, Information, and Technology*. Aspen Publishers, Inc., New York, NY.
- [138] Borce Stojkovski and Gabriele Lenzini. 2020. Evaluating ambiguity of privacy indicators in a secure email app. In *Italian Conference on Cyber Security*, Michele Loreti and Luca Spalazzi (Eds.). CEUR Workshop Proceedings, Vol. 2597. CEUR-WS Team, Aachen, Germany, 223–234.

- [139] Theeraporn Suphakul and Twittie Senivongse. 2017. Development of privacy design patterns based on privacy principles and UML. In *Proceedings of the 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD'17)*. IEEE Computer Society, Washington, DC, 369–375. <https://doi.org/10.1109/SNPD.2017.8022748>
- [140] Damian A. Tamburri. 2020. Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation. *Inf. Syst.* 91, Article 101469 (July 2020), 14 pages. <https://doi.org/10.1016/j.is.2019.101469>
- [141] Omer Tene and Jules Polonetsky. 2011. Privacy in the age of big data: A time for big decisions. *Stan. L. Rev. Online* 64 (Feb. 2011), 63–69.
- [142] Sigal Tifferet. 2019. Gender differences in privacy tendencies on social network sites: A meta-analysis. *Comput. Hum. Behav.* 93 (April 2019), 1–12. <https://doi.org/10.1016/j.chb.2018.11.046>
- [143] Shukun Tokas, Olaf Owe, and Toktam Ramezanifarkhani. 2020. Language-based mechanisms for privacy-by-design. In *Privacy and Identity Management. Data for Better Living: AI and Privacy*, Michael Friedewald, Melek Önen, Eva Lievens, Stephan Krenn, and Samuel Fricker (Eds.). IFIP Advances in Information and Communication Technology, Vol. 576. Springer Nature Switzerland AG, Cham, Switzerland, 142–158. https://doi.org/10.1007/978-3-030-42504-3_10
- [144] Janice Y. Tsai, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. 2011. The effect of online privacy information on purchasing behavior: An experimental study. *Inf. Syst. Res.* 22, 2 (February 2011), 254–268. <https://doi.org/10.1287/isre.1090.0260>
- [145] Bibi van den Berg and Simone van der Hof. 2012. What happens to my data? A novel approach to informing users of data processing practices. *First Monday* 17, 7 (July 2012), 15 pages. <https://doi.org/10.5210/fm.v17i7.4010>
- [146] †Yung S. Van Der Sype and Walid Maalej. 2014. On lawful disclosure of personal user data: What should app developers do? In *Proceedings of the 7th International Workshop on Requirements Engineering and Law (RELAW'14)*, Anna Perini, Daniel Amyot, Annie Antón, Travis D. Breaux, Aaron Massey, and Alberto Siena (Eds.). IEEE, New York, NY, 25–34. <https://doi.org/10.1109/RELAW.2014.6893479>
- [147] Max van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J. Weitzner, and Nigel Shadbolt. 2017. Better the devil you know: Exposing the data sharing practices of smartphone apps. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'17)*. ACM, New York, NY, 5208–5220. <https://doi.org/10.1145/3025453.3025556>
- [148] Evangelia Vanezi, Dimitrios Kouzapas, Georgia M. Kapitsaki, Theodora Costi, Alexandros Yeratziotis, Christos Metouris, Anna Philippou, and George A. Papadopoulos. 2019. GDPR compliance in the design of the INFORM e-Learning platform: A case study. In *Proceedings of the 13th International Conference on Research Challenges in Information Science (RCIS'19)*, Manuel Kolp, Jean Vanderdonckt, Monique Snoeck, and Yves Wautelet (Eds.). IEEE, New York, NY, 257–266. <https://doi.org/10.1109/RCIS.2019.8877022>
- [149] Konstantina Vemou and Maria Karyda. 2014. Guidelines and tools for incorporating privacy in social networking platforms. *IADIS Int. J. WWW/Internet* 12, 2 (2014), 16–33.
- [150] Isabel Wagner and David Eckhoff. 2018. Technical privacy metrics: A systematic survey. *ACM Comput. Surv.* 51, 3, Article 57 (July 2018), 38 pages. <https://doi.org/10.1145/3168389>
- [151] Huaqing Wang, Matthew K. O. Lee, and Chen Wang. 1998. Consumer privacy concerns about Internet marketing. *Commun. ACM* 41, 3 (3 1998), 63–70. <https://doi.org/10.1145/272287.272299>
- [152] Sarah Myers West. 2019. Data capitalism: Redefining the logics of surveillance and privacy. *Bus. Soc.* 58, 1 (July 2019), 20–41. <https://doi.org/10.1177/0007650317718185>
- [153] Alan F. Westin. 1967. *Privacy and Freedom*. Athenium, New York.
- [154] Alan F. Westin. 2003. Social and political dimensions of privacy. *J. Soc. Issues* 59, 2 (April 2003), 431–453. <https://doi.org/10.1111/1540-4560.00072>
- [155] Shomir Wilson, Florian Schaub, Aswarth Abhilash Dara, Frederick Liu, Sushain Cherivirala, Pedro Giovanni Leon, Mads Schaarup Andersen, Sebastian Zimmeck, Kanthashree Mysore Sathyendra, N. Cameron Russell, Thomas B. Norton, Eduard Hovy, Joel Reidenberg, and Norman Sadeh. 2016. The creation and analysis of a website privacy policy corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Erk Katrin and Noah A. Smith (Eds.). Association for Computational Linguistics, Berlin, Germany, 1330–1340. <https://doi.org/10.18653/v1/P16-1126>
- [156] Shomir Wilson, Florian Schaub, Frederick Liu, Kanthashree Mysore Sathyendra, Daniel Smullen, Sebastian Zimmeck, Rohan Ramanath, Peter Story, Fei Liu, Norman Sadeh, and Noah A. Smith. 2019. Analyzing privacy policies at scale: From crowdsourcing to automated annotations. *ACM Trans. Web* 13, 1, Article 1 (Feb. 2019), 29 pages. <https://doi.org/10.1145/3230665>
- [157] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE'14)*. ACM, New York, NY, Article 38, 10 pages. <https://doi.org/10.1145/2601248.2601268>

- [158] Gilbert Wondracek, Thorsten Holz, Engin Kirda, and Christopher Kruegel. 2010. A practical attack to de-anonymize social network users. In *Proceedings of the IEEE Symposium on Security and Privacy*. IEEE Computer Society, Los Alamitos, CA, 223–238. <https://doi.org/10.1109/SP.2010.21>
- [159] Heng Xu, Tamara Dinev, H. Jeff Smith, and Paul Hart. 2008. Examining the formation of individual’s privacy concerns: Toward an integrative view. In *Proceedings of the International Conference on Information Systems (ICIS’08)*. Association for Information Systems, Atlanta, GA, Article 6.
- [160] Le Yu, Xiapu Luo, Jiachi Chen, Hao Zhou, Tao Zhang, Henry Chang, and Hareton K. N. Leung. 2021. PPChecker: Towards accessing the trustworthiness of android apps’ privacy policies. *IEEE Trans. Softw. Eng.* 47, 2 (2021), 221–242. <https://doi.org/10.1109/TSE.2018.2886875>
- [161] Le Yu, Tao Zhang, Xiapu Luo, and Lei Xue. 2015. Autoppg: Towards automatic generation of privacy policy for android applications. In *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM’15)*. ACM, New York, NY, 39–50. <https://doi.org/10.1145/2808117.2808125>
- [162] Sebastin Zimmeck, Rafael Goldstein, and David Baraka. 2021. Privacyflash pro: Automating privacy policy generation for mobile apps. In *Proceedings of the 28th Network and Distributed System Security Symposium (NDSS’21)*. The Internet Society, Reston, Virginia, 18 pages. <https://doi.org/10.14722/ndss.2021.24100>
- [163] *Steven Zimmerman, Alistair Thorpe, Chris Fox, and Udo Kruschwitz. 2019. Investigating the interplay between searchers’ privacy concerns and their search behavior. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR’19)*. ACM, New York, NY, 953–956. <https://doi.org/10.1145/3331184.3331280>
- [164] Sergio Zorzo, Diego Pontes, Jose Mello, and Diego Dias. 2016. Privacy rules: Approach in the label or textual format. In *Proceedings of the 22nd Americas Conference on Information Systems: Surfing the IT Innovation Wave (AMCIS’16)*. Association for Information Systems, Atlanta, GA, 10 pages.
- [165] Tomasz Zukowski and Irwin Brown. 2007. Examining the influence of demographic factors on internet users’ information privacy concerns. In *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries (SAICSIT’07)*. ACM, New York, NY, 197–204. <https://doi.org/10.1145/1292491.1292514>

Received September 2020; revised November 2021; accepted November 2021