Computers & Security 146 (2024) 104065

Contents lists available at ScienceDirect

Computers & Security

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

Privacy policy analysis: A scoping review and research agenda

Karl van der Schyff [a,*], Suzanne Prior [a], Karen Renaud [a,b]

[a] School of Design and Informatics, Abertay University, Dundee, UK
[b] Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK

ARTICLE INFO

Keywords:
Privacy policy analysis
Privacy policy classification
Privacy policy benchmarking
Privacy policy completeness
Privacy policy rule
Privacy policy strategy
Machine learning
Privacy policy evaluation
Scoping review

ABSTRACT

Online users often neglect the importance of privacy policies - a critical aspect of digital privacy and data protection. This scoping review addresses this oversight by delving into privacy policy analysis, aiming to establish a comprehensive research agenda. The study's objective was to explore the analytic techniques employed in privacy policy analysis and to identify the associated challenges. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews (PRISMA-ScR) checklist, the review selected $n = 97$ relevant studies. The findings reveal a diverse array of techniques used, encompassing automated machine learning and natural language processing, and manual content analysis. Notably, researchers grapple with challenges like linguistic nuances, ambiguity, and complex data harvesting methods. Additionally, the lack of privacy-centric theoretical frameworks and a dearth of user evaluations in many studies limit their real-world applicability. The review concludes by proposing a set of research recommendations to shape the future research agenda in privacy policy analysis.

1. Introduction

Privacy is a critical concern for individuals and organizations across the globe (Mohammadi et al., 2019; Solove, 2021). As technology advances and data collection becomes pervasive, the design of privacy policies has gained significant importance (Ghazinour and Albalawi, 2016). This stems from privacy policies, in effect, serving as legal documents outlining how organizations handle personal information. Importantly, privacy policies contain increasingly complex (and lengthy) legal content (Alabduljabbar et al., 2021a; Meier et al., 2020). This stems from intensified privacy concerns, which necessitated the development of lengthier and more complex policies to address potential legal risks. For example, legal frameworks, such as the General Data Protection Regulation (GDPR), impose strict obligations on organizations, which, in turn, further adds to the complexity of modern privacy policies. Unfortunately, this results in decreased motivation to read and understand said privacy policies (Acquisti et al., 2020). In fact, research shows that individuals spend only a few seconds or minutes scanning privacy policies due to time constraints, cognitive overload, and a lack of trust in organizations' transparency (Obar and Oeldorf-Hirsch, 2020). This confound is further compounded by convoluted sentence structures used by many privacy policies (Neal et al., 2023). Together, the above leads to a situation that perpetuates:

- Users unwittingly consenting to data practices they would usually be uncomfortable accepting (Boliek, 2021; Liao et al., 2020; Shayegh et al., 2019).
- An increased sense of mistrust towards organizations, perceiving privacy policies as deceptive or designed to obscure rather than inform (Kretschmer et al., 2021; Zhu et al., 2020).
- Privacy-conscious individuals actively disengaging from digital services (Acquisti et al., 2020; Lau et al., 2018).

Given the importance of privacy policies, much research has been done to streamline or optimize them. For example, recent research has focused on:

- Simplifying privacy policies to enhance user comprehension and engagement by using plain language, clear structure, and concise explanations of data practices (Mohammadi et al., 2019; Sanghavi et al., 2022).
- Adopting a layered approach to privacy policy design, where essential information is presented upfront, and additional details are provided in separate sections, enabling users to access information based on their specific needs and interests (Gerl and Meier, 2019; Leicht et al., 2021).

* Corresponding author.
E-mail address: k.vanderschyff@abertay.ac.uk (K. van der Schyff).

https://doi.org/10.1016/j.cose.2024.104065
Received 25 March 2024; Received in revised form 3 June 2024; Accepted 15 August 2024
Available online 18 August 2024

- Incorporating visual elements, such as infographics or icons, to aid in conveying complex concepts and data flows in a more accessible and engaging manner (Efroni et al., 2019), and
- Promoting privacy literacy through educational initiatives to empower individuals to make informed decisions and better understand the implications of their data disclosures (Ebert et al., 2021; Soumelidou and Tsohou, 2020).

Despite the plethora of research on privacy policies, there is a distinct lack of studies that systematically synthesize existing work on privacy policy analysis in a platform-agnostic manner. For example, while the review by Del Alamo et al. (2021) provides researchers with a comprehensive summary of privacy policy analysis and assessment techniques, their work focuses solely on Android apps. The review assessed studies published between 2016 and 2020, and updated reviews could deliver new insights. Similarly, Del Alamo et al. (2022) focused on studies that performed automated text-based privacy policy analyses using Natural Language Processing (NLP).

The review by Bhattacharjee et al. (2020) focused on the visual aspects of privacy policy analysis - particularly how visualizations are used as privacy-preserving technologies. Leicht and Heisel (2019), on the other hand, focused on privacy policy languages – an area that has undergone significant change since 2012 as many organizations no longer use privacy languages (e.g., P3P). Although conceptually like our review, Liu et al. (2022) focused on machine learning and its role as a privacy-preserving technology. To date, many of the reviews have therefore been very *narrow*; summarizing extant privacy policy analysis research for specific audiences using *one technique or platform*.

A broader investigation of the privacy policy analysis research landscape is needed and could help to deliver valuable insights. As such our rationale for this study stems from the recognition that there is a need to explore the various privacy policy analysis techniques and their associated challenges broadly. Furthermore, the ambiguity, linguistic nuances, and complex nature of privacy policy content necessitate a more nuanced and interdisciplinary approach to analysis. Our review contributes on three fronts. First, we provide a comprehensive overview of 97 studies, offering insights into the various analytic techniques used in privacy policy research, from machine learning and natural language processing to manual content analysis. Second, we identify the primary challenges researchers face, such as linguistic ambiguity, policy harvesting complexities, and the lack of privacy-centric theoretical frameworks. Third, we propose a detailed research agenda with theoretical and practical recommendations to guide future research in this domain. By emphasizing the integration of user evaluations and the application of established privacy theories, we aim to enhance the real-world applicability and impact of privacy policy research. To assist in this regard, we posed the following research questions:

RQ1. Which analytic techniques have been used when performing privacy policy analyses? This question's purpose is to understand the variety of techniques available to future researchers in this field. It would be beneficial to correlate classifications of these techniques with the other aspects of this review, such as the corresponding challenges we identify.

RQ2. What challenges have deterred privacy policy analyses? This question is to summarize the core challenges a researcher may face when performing privacy policy analysis. These will align themselves with the specific techniques identified, which is beneficial, enabling a researcher to make an informed decision when selecting an analysis technique.

We argue that this investigation will help researchers to select the most appropriate analysis technique for the context at hand. Moreover, they will have an idea of the challenges its use could entail as well as the likely outcomes based on previous empirical work. Our review is structured as follows: The methodology of the scoping review is described in Section 2 below, including an overview of the data charting process. This is followed by a summary of the findings in Section 3. Section 4 presents readers with a discussion of our proposed research agenda. The study's limitations and areas of future research are outlined in Section 5, followed by a conclusion in Section 6.

2. Methodological approach

A thorough and systematic approach to selecting articles is essential to ensure the inclusion of relevant material while conducting a scoping review. Therefore, several steps are used in our search selection procedure to find and retrieve articles which may be appropriate for this study. The search selection techniques used in our scoping review are described in the following Sections. In particular, we used Arksey and O'Malley's scoping review methodology to formulate our search string. The use of a scoping review (as opposed to a systematic review) is methodologically appropriate, because:

- Scoping reviews are an ideal way to explore complex research areas comprised of many concepts, techniques, and approaches. For example, within this study, we wished to "examine" how others have researched privacy policy analysis (Munn et al., 2018; Peters et al., 2015).
- A scoping review is useful to outline research areas that will likely yield useful results. We find this to be particularly useful in an interdisciplinary context where it's not easy to see how the concepts within a research landscape *are used together* (Munn et al., 2018). This aligns well with the research agenda we developed in this study.
- Scoping reviews enables researchers to include a wide range of study designs *from a variety* of academic outlets. This, in turn, makes them uniquely suited to areas where researchers wish to get a better understanding of "what has been done thus far" irrespective of specific quality assessments that typify systematic reviews (Kitchenham et al., 2011). This makes scoping reviews ideal for exploring complex and heterogeneous fields where different study designs, methodologies, and theoretical frameworks are typically employed - as is the case with privacy policy analysis. Arksey and O'Malley (2005) are explicit in this regard stating that scoping reviews are valuable for examining the extent, range, and nature of research activities *where the evidence is diverse*.
- Scoping reviews serves as a preliminary assessment of the volume and nature of research available on a topic. In other words, they assist researchers in determining the feasibility and scope of a future systematic review (Munn et al., 2018; Peters et al., 2015). In this regard our research agenda serves as a useful point of departure for future systematic reviews where researchers may wish to focus on *extremely specific problems*.

We report the results of our study selection procedure using the Preferred Reporting Items for Systematic Reviews extension for Scoping Reviews (PRISMA-ScR) checklist for scoping reviews (Tricco et al., 2018) in the following section.

2.1. Search strategy

When conducting a review, a comprehensive search strategy should be devised to cover relevant academic databases within the field under consideration. In this regard, and because this study spans both information systems and computer science, we also searched within the A* and A-rated Information Systems journals ($n = 56$) that appear on the 2022 version of the Australian Business Deans Council (ABDC) list[1]. Although additional sources, such as grey literature and relevant organizational websites, may also be explored to minimize publication bias,

[1] https://abdc.edu.au/abdc-journal-quality-list/

we opted to focus on peer-reviewed material. Our motivation for taking such a hybrid approach stems from the fact that we wanted to conduct an interdisciplinary review that would include behavioral, legal, and technological articles. After obtaining ethical clearance (ref EMS7267), we searched within ten academic databases, including *Sage, IEEE Xplore, ACM, ScienceDirect, Emerald Insight, Springer, Taylor & Francis, Google Scholar,* and *Wiley*. A combination of regulated vocabulary items (such as Medical Subject Headings - MeSH terms) and pertinent keywords should be employed to enhance the search process. Additionally, it is crucial that the search phrases be directly aligned with the review's research questions and objectives. Using the above guidelines, we formulated the following search string:

"privacy policy classification" OR "privacy policy benchmarking" OR "privacy policy completeness" OR "privacy policy evaluation" OR "privacy policy analysis" OR "privacy policy integrity" OR "privacy policy assessment" OR "privacy policy test" OR "privacy policy visualization" OR "privacy policy visualisation"

Note that specific search phrases that expressly refer to a particular user or type of context were excluded. This was done purposefully, given our second review question, which aims to understand the challenges facing all policy analysis contexts. Search results were exported in either BibTeX, Comma Separated Value (CSV), or Research Information Systems (RIS) formats.

2.2. Eligibility criteria

After obtaining initial search results, various screening processes commenced. Studies were selected for detailed review if they satisfied the following criteria:

- Results or findings were based on the analysis of empirical data.
- A focus on online privacy policy analysis, as defined in the introduction.
- The studies were published between 2004 and 2023. We chose 2004 as it coincided with the emergence of Web 2.0 (Cooke and Buckley, 2008; John, 2013).
- A focus on any form of online user or system. This included the internet, social media, web apps, websites, and government systems.

Studies were excluded if they matched the following criteria:

- They were mostly conceptual, containing little (or no) empirical support for their findings.
- Scoping, systematic, or meta-analytic reviews.
- Addressed privacy policy analysis outside of a technological setting.
- Not written in English.
- Not peer-reviewed.

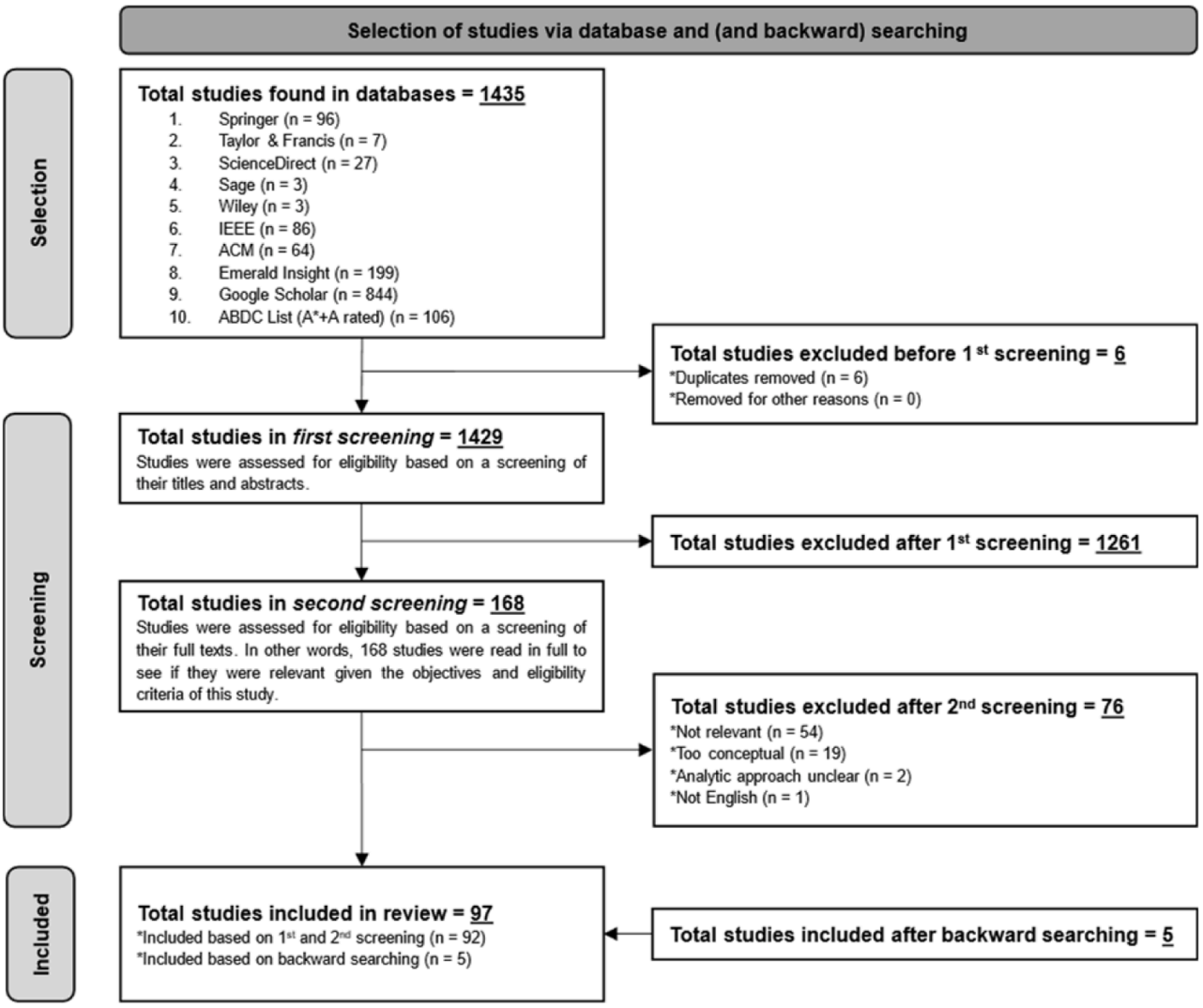

Fig. 1. Visual summary of our PRISMA-ScR flow diagram.

2.3. Study selection process

We initially screened the titles and abstracts using the inclusion and exclusion criteria during the study selection process. This stage involved all the authors to guarantee consistency and reduce bias. Studies that failed to satisfy the eligibility requirements were excluded. Conversely, studies with some uncertainty or potential for relevance moved on to the 2nd stage of the selection process (Levac et al., 2010). During the full-text screening, each article was thoroughly examined by the first two authors considering the inclusion and exclusion criteria (Pham et al., 2014). This entire process is visually summarized in Fig. 1.

2.4. Data charting and analysis

After completing the full-text screening, relevant data was retrieved to populate several data extraction tables. These often focus on a study's general characteristics (e.g., methodology used or empirical situations), significant findings, and other pertinent information aligned with the research questions. Typically, one or more authors extract the data, and disagreements are settled by consensus or discussion with the study team (Tricco et al., 2018). In this instance, we resolved any disputes by consulting with the third author. The data extraction tables of this review contain analytic information aligned to the following:

- **The characteristics of the studies under review.** This includes their *privacy context, country of study, study design, theoretical framework, datasets (new and existing),* and *participants used*. It is worth noting that our analysis revealed that many studies made use of existing datasets. As such, we subdivided this category into the use of existing empirical data and those that generated their own data. These were recorded as *existing* and *new* within the data extraction table.
- **The analytic techniques used.** These have been categorized as far as possible. Still, it is worth noting that the prevalence of Machine Learning (ML) led us to analyze the resultant data by performing two co-occurrence analyses. The first focused on the use of techniques that co-occur with the use of ML, and the other focused on those techniques that overlap when ML is not used. Some examples of the latter include *visualization, word counts, readability testing, and statistical modeling* (amongst others). Because the field of machine learning is complex with many variations, we did not attempt to subdivide the techniques further.
- **The challenges encountered when applying the associated policy analysis techniques.** Our analysis also reports on the related effectiveness of the methods used (both technical and user-focused). Specifically, within the context of the *findings*, and the *challenges* identified. We used the latter (and their associated research implications) to develop the research agenda, but avoided further interpretation of effectiveness as this can be subjective.

See Tables A.1 through A.5 in the appendix for a complete outline of the data extraction tables associated with the above.

3. Findings

After removing duplicates, the study selection procedure yielded 1429 studies. This was reduced to 168 after screening the titles and abstracts (1st screening stage). We further reduced this number by performing a full-text screening (2nd stage) of these studies, leading to the inclusion of 97 studies in the final review.

3.1. Overview of selected studies

Most of our review studies were conducted in the United States (n = 41), followed by Germany (n = 8), the United Kingdom (n = 8), China (n = 7), and Canada (n = 5). We noted an apparent absence of studies originating in Africa, with very few originating from Australasia (n = 3). From a contextual point of view, most studies focused on analyzing website privacy policies (n = 45). Several (n = 8) used Alexa to select which websites to focus on (e.g., top 1000 websites based on traffic generated). A substantial portion of the remainder of the studies focused on mobile apps (n = 36). Regarding the empirical data used (or generated), we found that most studies (n = 72) avoided using existing policy datasets, opting to create their own. It is worth noting that those who did use existing policy datasets primarily used the OPP-115 dataset (n = 14) (Usableprivacy.org, 2017). Only two studies used the APP-350 (Zimmeck et al., 2019) dataset, focusing on Android app privacy policies. From a study design perspective, we found most policy analysis research to be either purely quantitative (n = 50) or mixed (quantitative and qualitative) (n = 27). Only eight studies used a strictly qualitative study design, and very few used established theoretical frameworks (n = 5). In fact, of the latter five, only three were privacy specific. Only one study was unclear about their methodological approach (Vanezi et al., 2021).

3.2. Analytic techniques used

RQ1. Which analytic techniques have been used when performing privacy policy analyses? Although we encountered various analytic techniques, most of the studies used machine learning (n = 61). Based on this finding, we set out to understand how the other deployed techniques relate to (or co-occur) with the use of machine learning (or not). As stated earlier, we created two categories – one to group studies where a machine is used and another where it is not used. For example, some studies used traditional machine learning techniques (e.g., Support Vector Machine (SVM) etc.), combined these techniques with natural language processing, or only performed natural language processing. Given the complexity of this field (i.e., ML) and the various techniques one can use, these studies were all classified as machine-learning based. Moreover, because some studies used more than two techniques, we could not use two-dimensional visualizations such as radial network diagrams. Instead, we developed co-occurrence matrices (see the appendix). Such matrices are particularly useful when highlighting the relationships between more than two dimensions in a complex system where co-occurrence is likely.

3.2.1. Absence of machine learning

Our findings indicate that in the absence of machine learning, manual analysis techniques (labeled as *Manual* in Table A.2) are most prevalent (n = 22). Five studies exclusively used manual analysis techniques (Bachiri et al., 2018; Bookert et al., 2022; Novikova et al., 2020; West, 2022; Zhao et al., 2020). See Table 1 for a definition of these techniques.

To derive meaningful recommendations, we investigated the level of co-occurrence amongst all the techniques used in the absence of machine learning. Although the manual analysis techniques dominate this category, they are often accompanied by visualizations (n = 8). Two of the latter studies used additional analytic methods, including word counts (Akanfe et al., 2020), in addition to statistical modeling and readability tests (but excl. readability indexes) (Reeder et al., 2008). The use of variance analyses (e.g., ANOVAs) seemed instrumental in these instances as researchers were able to measure the variance in understanding between various visual designs statistically. In addition to visualizations, the use of word counts, and readability analyses were found to co-occur with manual analyses in seven studies, with one of these only using word counts (Akanfe et al., 2020).

The readability indices we encountered varied, but most used the *Flesch-Kincaid* or *Gunning-Fog* readability indexes. From a statistical point of view, and in the case of two annotation-based studies, researchers made use of *Mann-Whitney-U* (also known as *Mann-Whitney-Wilcoxon*) to perform comparisons between the privacy policies of Android mobile money services with traditional US banks (Bowers et al.,

Table 1
Definition of manual techniques used to analyze privacy policies.

Technique	#	Summary	Refs.
Annotation	8	Given the legal context of privacy policies, several studies used expert annotators to manually annotate (and thus classify) the privacy policy statements. These studies often did so by referring to an accepted regional privacy policy guideline (e.g., GDPR). Often, the annotators were either fully trained legal professionals or students of a legal faculty. In one instance, the process of annotation was referred to as tagging.	Bookert et al. (2022), Bowers et al. (2019), Javed et al. (2021), Kaplan et al. (2021), Shvartzshnaider et al. (2019), Sunyaev et al. (2015), West (2022), Zhao et al. (2020)
Analysis of Questions	1	Only one study used this technique to develop a virtual assistant privacy awareness dashboard where questions pertaining to each privacy subject area were analyzed.	Bolton et al. (2023)
Category Assignments	2	Although like annotation, we felt these two studies should be listed separately as the categories used contained more concepts than those prescribed by the GDPR. In the case of , the privacy policies of WhatsApp and RestAssured were manually analyzed as case studies by developing a privacy policy template from the resultant qualitative analysis.	Akanfe et al. (2020), Mohammadi et al. (2019)
Content Analysis	6	These studies performed content analysis. Given the absence of theoretical frameworks, we argue that these content analyses were not guided by theory, which relates directly to our first recommendation within our research agenda.	Bachiri et al. (2018), Bareh (2022), Earp et al. (2005), Fox et al. (2018), Kandil et al. (2018), Yuan et al. (2023)
Heuristic Evaluation	1	Interestingly, only one study performed a heuristic evaluation, which used the privacy policy visualization model to evaluate Facebook's privacy policy.	Ghazinour and Albalawi (2016)
Interpretation (readability)	1	This study specifically mentions interpretation as a means of analysis, which also focuses on measuring the readability of 113 organizations' privacy policies.	McRobb and Rogerson (2004)
Layered Design	1	Although several other techniques are used, the researchers of this study manually developed a layered design to convey the privacy policy of six e-commerce websites. The layered design was contrasted with another format they developed in this study called privacy finder.	Reeder et al. (2008)
Ontology Assignment	1	An ontology is developed within the context of IoT privacy policies to perform a privacy risk assessment.	Novikova et al. (2020)
Policy Assessment	1	The authors use a policy assessment technique to complement their readability analysis of the privacy policies of 94 IoT devices. Having said this, there is no mention of the specifics of this privacy risk assessment (see recommendation two about privacy risk assessments).	Paul et al. (2018)

2017) as well as middle eastern banks (Javed et al., 2021). Interestingly, two of the five studies that did use a theoretical framework (i.e., *Contextual Integrity*) did so when using annotation on social media websites' privacy policies (Kaplan et al., 2021; Shvartzshnaider et al., 2019). See Table A.2 in the appendix for a complete outline of this co-occurrence matrix.

3.2.2. Presence of machine learning

Overall, 61 studies used machine learning, often in combination with several other analysis techniques. We found that annotation usually ($n = 22$) co-occurred with machine learning and to a lesser extent with word counts ($n = 5$), readability tests ($n = 4$), visualizations ($n = 4$), Hybridized Task Recomposition (HTR) ($n = 1$), and additional statistical modeling ($n = 1$). Although six studies mentioned crowdsourcing, we do not view this as an analysis technique but rather as a means of data collection. At least two studies made use of Amazon's Mechanical Turk (Liu et al., 2014; Qiu and Lie, 2020), with others like Kotal et al. (2021) recruiting general web users as annotators. When machine learning occurred with readability testing, these studies used the *Flesch-Kincaid* (Libert, 2018; Zimmeck and Bellovin, 2014) index exclusively or calculated an average of a privacy policies' readability by using the *Flesch-Kincaid, Gunning Fog,* and *SMOG* indices (Wettlaufer and Simo, 2020). One study focused on Chinese privacy policies and used complex distance calculations to calculate a readability score (Lin et al., 2022). Few studies focused on analyzing non-English privacy policies (see recommendations).

Bhatia et al.'s (2016) use of HTR was an interesting and novel approach to the use of crowdsourced annotation, striking a balance between cost, complexity, and effectiveness. A similar novel use of Device Attribute Mapping (DAM) was used by Manandhar et al. (2022) to perform a comparative study of 2442 smart home device privacy policies. Only one study used statistical modeling in addition to machine learning—specifically, *Mann-Whitney-Wilcoxon* testing and Covariance-Based Structural Equation Modeling (CB-SEM). Interestingly, none of the studies we reviewed used Partial Least Squares Structural Equation Modeling (PLS-SEM). This is interesting as PLS-SEM is well suited to developing composite models where one could empirically evaluate an artifact or visualization. See Table A.3 in the appendix for a complete outline of this co-occurrence matrix.

3.3. Challenges

RQ2. What challenges have deterred privacy policy analyses? To extract the relevant challenges, we thematically analyzed the data in Tables A.4 and A.5 (specifically the *Challenges* column). Significantly, and to correlate the challenges to specific sets of techniques (i.e., those that co-occur), we thematically analyzed the challenges within studies that did not use ML (as well as those that did). This enabled a comparison of thematic findings with our earlier analysis. Note that some studies did not clearly articulate the challenges encountered (Alabduljabbar et al., 2021b; Alshamsan and Chaudhry, 2022).

3.3.1. Challenges: machine learning studies

We identified five themes related to the challenges researchers may encounter when using machine learning as a core privacy policy analysis technique.

3.3.1.1. Analytic complexity. Our analysis indicates that several factors increase the complexity of ML-based privacy policy analyses (Al Rahat et al., 2022; Adhikari et al., 2022; Chang et al., 2020; Costante et al., 2012; Guamán et al., 2023). Prominent among these are the influence of ambiguity and privacy policy harvesting (Guntamukkala et al., 2015). Ambiguity is mentioned by the developers of *PolicyLint* – a privacy policy analysis tool whose machine learning results need to be manually verified by researchers to combat ambiguity (Aberkane et al., 2022; Andow et al., 2019). Similar issues are reported by Sarne (2019), Story et al. (2019), who explicitly states that expert-based post-processing is often required – especially when not performing topic modeling. Within this context, ambiguity (and related complexity) refers to those instances where annotated policy text could be interpreted in many ways, complicating labeling within the dataset used by the ML algorithm.

Additional issues are reported by Bateni et al. (2022), who report problems related to the ambiguity inherent in the seed words used in their ML analyses, with several other studies also reporting annotation-related ambiguity (Hashmi et al., 2022; Mousavi Nejad et al., 2020; Oltramari et al., 2018; Subahi and Theodorakopoulos, 2023). Ambiguity also plagues studies focused on developing privacy policy ontologies, as the related ontological models are subjective by nature (Audich et al., 2021; Oltramari et al., 2018).

Several studies explicitly mention the complexities inherent in policy harvesting. Policy harvesting seems particularly challenging when performing longitudinal privacy policy analysis (Hashmi et al., 2022). For example, Amos et al. (2021) found policy harvesting to be complex despite focusing on popular websites with highly structured policies. Similar challenges were reported by Hashmi et al. (2022), who investigated the extent to which apps leak data and how that correlates with the harvested policies over several years pre-GDPR. Policy harvesting is also reported to complicate policy analyses, which compare network traffic with the information the policy explicitly states will be tracked (Bui et al., 2022). Similar challenges are encountered when analyzing the privacy policies of Android mobile apps (Bui et al., 2021) and OVR apps (Trimananda et al., 2022).

3.3.1.2. Language and semantic variation. In addition to ambiguity and analytic complexity, several studies hinted at the additional challenges posed by different languages and semantic structures. The latter challenges introduce an extra layer of intricacy that can be time-consuming and challenging when performing privacy policy analyses (Ahmad et al., 2020); Asif et al., 2021; Nokhbeh Zaeem et al., 2020). This is further compounded by challenges related to semantic dependencies and limitations as to the lexicons in use (Ahmad et al., 2022; Audich et al., 2018; Bhatia et al., 2016). Bracamonte et al. (2019) found this particularly challenging when devising a means to communicate the risk within the context of Japanese privacy policies.

As stated, semantic structures also pose significant challenges. More so in terms of deciding on the optimal semantic unit of analysis. For example, Kotal et al. (2021) and Liu et al. (2023) found it challenging to attain accurate results when classifying text at the paragraph level. This is particularly true when there are variations in language use, which inevitably increase the level of ambiguity. The results reported by Liepin et al. (2019) and Yang et al. (2021) further complicate matters. They found that focusing on single sentences makes extracting information about unfair or unlawful data use and rule development challenging. Additionally, many tools to analyze privacy policies only support English privacy policies (Tesfay et al., 2018; Wagner, 2023). To overcome these challenges, Liepin et al. (2019) suggest that supervised ML should be combined with techniques centered around neural artificial intelligence (see recommendations). More fundamentally, when the policies originate from foreign websites, translation becomes indispensable but often fails to capture the nuanced meaning of the original language (Lin et al., 2022; Liu et al., 2022). Keyword extraction, without context in some cases (Mousavi Nejad et al., 2018), poses yet another challenge – especially if used within unsupervised contexts (see recommendations) or where static analysis techniques are used (Yu et al., 2016).

3.3.1.3. Visualization and user evaluation. Although only a few studies used visualizations in combination with ML, we identified two challenges plaguing these studies. The first relates to the lack of user evaluation when studying various privacy designs (Cui et al., 2023). One example is the study by Dombetzki et al. (2020), which focuses heavily on the design process of their privacy plugin (Amaryllis) but avoids complementing their use of machine learning with a user evaluation. We argue that this severely limits the applicability of such privacy artifacts (see recommendations). This ties into the second challenge, which relates to participant confusion when studying visual representations of privacy policies. It is clear from the study by Reeder et al. (2009) that interactive grids are particularly susceptible to this, even more so when performing a direct comparison with older forms of privacy grid designs such as P3P, which is unused.

3.3.1.4. Data collection. During our analysis, several data collection related challenges emerged, including the scarcity of privacy policies in languages other than English, the absence of policies (especially mobile apps), and the pre-GDPR nature of some datasets (Hamdani et al., 2021). Additionally, the reliance on the Wayback Machine, as well as the OPP-115 data corpus, introduces challenges in capturing policies comprehensively, leading to potential omissions that limit the statistical inferences that can be made (Harkous et al., 2018; Khandelwal et al., 2023; Ahmad et al., 2020). Similar challenges were evident among studies that employed specific sampling strategies, focused only on types of apps (Hatamian et al., 2019, 2021), or used unknown application programming interfaces – all of which may lead to the possibility of biased results (Bui et al., 2022; Chang et al., 2019). Crowdsourcing introduces challenges, particularly concerning non-trivial annotations by crowd workers who do not always find it easy to accurately perform annotations (Wilson et al., 2016; Zimmeck and Bellovin, 2014).

The potential for interpretive variations among annotators and unequal commitment to the project (Liu et al., 2014) further challenge researchers who try to attain high data quality and consistency in their analyses (Chaw and Chua, 2021). Additionally, the emotional charge often associated with user reviews (Wettlaufer and Simo, 2020) can complicate the investigation, potentially biasing results toward privacy aspects that elicit strong emotions. These issues must be accounted for when collecting data – *especially those used for training.* Although previously mentioned, the ambiguity within specific policies also introduces complexities in data collection when working with large datasets (Thotawaththa et al., 2021). For example, should the data be collected at the sentence or paragraph level? We have already established that such decisions lead to further challenges in policy analysis.

3.3.1.5. Temporal nature of legal and policy terms. A few studies mentioned challenges related to the evolution of terms and legal definitions. For example, Narksenee and Sripanidkulchai (2019) explicitly state that ranking classification algorithms without clear definitions to guide classification is challenging. This is different from ambiguity because, although some studies are clear regarding the definition of privacy disclosure, this may change over time, impacting comparative (and longitudinal) work. Challenges related to the temporal nature of these legal terms and definitions are also mentioned by Nokhbeh Zaeem and Barber (2021) and, to a lesser extent, by Yu et al. (2016), who also encountered sentence-level limitations when using static ML techniques

(see recommendations).

3.3.2. Challenges: non-machine learning studies

We identified three themes related to the challenges researchers may encounter when they analyze privacy policies without the assistance of machine learning.

3.3.2.1. Analytic limitations and subjectivity. Without machine learning, clear limitations apply relating to the subjectivity involved in content analysis (Bareh, 2022; Earp et al., 2005; Ghazinour and Albalawi, 2016; Javed et al., 2021; McRobb and Rogerson, 2004) and their complex interpretive nature (Bachiri et al., 2018; Paul et al., 2018). As is the case for machine learning studies, these limitations revolve around the complexity (Bowers et al., 2019) and ambiguity of the data being analyzed – especially when using crowd workers (Shvartzshnaider et al., 2019). This situation is further complicated by the varied methods employed, which collectively introduce subjectivity and ambiguity into the analysis. Similarly, several studies performed readability analyses but used a wide range of techniques to make similar statements about the complexity of a privacy policy. Having said this, not all the readability indexes test precisely the same aspect of a policy – a concept that not all studies made clear (Becher and Benoliel, 2021; Bowers et al., 2019; Reeder et al., 2009).

It appears as if readability analyses, although complex, are sometimes treated as silver bullets[2] when it comes to complexity analysis (Bowers et al., 2019; Cadogan, 2011). Some non-ML studies also used calculations derived from unknown instruments, making it challenging to compare results as the validity and reliability of these are indeterminate within the larger privacy policy research community (Kaplan et al., 2021). Additionally, although user studies within non-ML policy research are the norm, some provided little feedback, limiting the applicability of the results (Tucker et al., 2015).

3.3.2.2. Data scope and collection. Another challenge we encountered was that of using datasets with limited scope. Several studies focused on specific subsets of data, such as particular industries (Bowers et al., 2019; Cottrill and Thakuriah, 2011; Jilka et al., 2021), demographics (Drozd and Kirrane, 2020), or geographic regions (Bookert et al., 2022) – specifically the US. This limited scope makes it challenging to generalize (Kandil et al., 2018) and apply the results to other contexts where privacy policies are either absent (Sunyaev et al., 2015) or the terms used are vaguely described (Bolton et al., 2023; West, 2022), and thus difficult to use definitively (and comprehensibly). In some instances, a limited set of applications or app stores were used (Farooq et al., 2020; Paspatis et al., 2020) with two qualitative studies making use of only one case study (Novikova et al., 2020; Zhao et al., 2020). Fox et al.'s (2018, 2022) studies also clarify that researchers should be wary when collecting data using visual study aids (e.g., nutrition labels). In such instances, unless correctly designed, the experiments could inadvertently prime participants.

3.3.2.3. Visual and technical challenges. Our analysis also points to several challenges related to the visual representation of data and the technical aspects of analysis, such as unclear software choices (Akanfe et al., 2020). For example, Yuan et al. (2023) state that it is challenging to illustrate real-world scenarios when studying third-party data flows visually. Similarly, Mohammadi et al., 2019 and Guo et al. (2020) state that even if a sufficient visual design could be developed, users are often not familiar enough with the icons, making it challenging to interpret the results. The same applies to the research conducted by Kelley et al. (2009, 2010, 2013) and McDonald et al. (2009) where it is evident that participants find it challenging to interpret the meaning of nutrition label text and design choices (e.g., blank spaces and assorted colors). For example, participants have vastly different understandings of what the phrase personal information means. Some may argue that such cases require additional pre-experiment training, which may verge on priming.

[2] Terminology from The Mythical Man Month by Fredrick P. Brooks (1995)

4. Discussion and research agenda

The overarching objective of this scoping review was to develop a research agenda by addressing two research questions focused on understanding which privacy policy analysis techniques are used, and what challenges researchers have encountered when using these techniques. Below, we discuss the implications of our findings by making a series of research recommendations, which, together, comprise our research agenda. As before, we separate our recommendations and provide a plan for ML-based approaches and those that do not use ML approaches. We commence our discussion with several general research recommendations that apply to any research focused on the analysis of privacy policies.

4.1. General guidance for privacy policy research

Before moving on to technique-specific recommendations, we want to convey some general research recommendations that should be noted as foundational. These apply regardless of the use of machine learning. We argue that these will improve the overarching quality and approach of a study focused on the analysis of privacy policies. Fig. 2 provides a complete overview of the recommendations provided in this section.

Recommendation 1: *To enhance the rigor and effectiveness of privacy policy analysis research, it is highly recommended that researchers* ***make extensive use of established privacy-centric theoretical frameworks****.* Established theoretical frameworks provide a solid foundation for defining and understanding key privacy concepts. These frameworks offer a structured vocabulary and precise definitions that enable researchers to communicate effectively, reducing ambiguity and ensuring consistent terminology in their analyses. Theories such as *Contextual Integrity* (Nissenbaum, 2019; Shvartzshnaider et al., 2019), *Privacy Calculus* (Dinev et al., 2006; Laufer and Wolfe, 1977), or the *Concept of Informational Self-determination* (Buitelaar, 2017) offer a holistic perspective on privacy issues. By embracing these frameworks, researchers gain a more comprehensive understanding of the multifaceted nature of privacy, allowing for a more nuanced analysis of policy implications. This is crucial as our findings indicate that policy research is susceptible to varied interpretations and ambiguity. We argue that a strict focus on a core set of constructs will reduce such ambiguity.

Additionally, privacy research is inherently interdisciplinary, spanning law, ethics, psychology, sociology, and technology. Established privacy-centric frameworks bridge these disciplines, fostering a more inclusive and integrated approach to research. Considering the broader societal implications of privacy policies, this interdisciplinary perspective is crucial for policy analysis. Using consistent theoretical frameworks also enables researchers to make meaningful comparative analyses. Such a comparative approach allows for benchmarking and identifying best practices, facilitating the development of more effective and universally applicable privacy policies. Our findings indicate that much of the privacy policy analysis research avoids the user. This is a critical omission which should be rectified going forward. Performing blanket GDPR-based compliance checking is important, but these systems are used by individuals, requiring us to evaluate the actual impact on individuals and society. Privacy frameworks can assist in this regard as they often include elements that assess the potential consequences of policies on individuals' autonomy, trust, and rights, offering a more comprehensive perspective on their implications.

Recommendation 2: *Privacy policy analysis research should* ***incorporate established privacy risk assessment methods*** *and* ***include a comprehensive Data Protection Impact Assessment*** *(DPIA).* Established

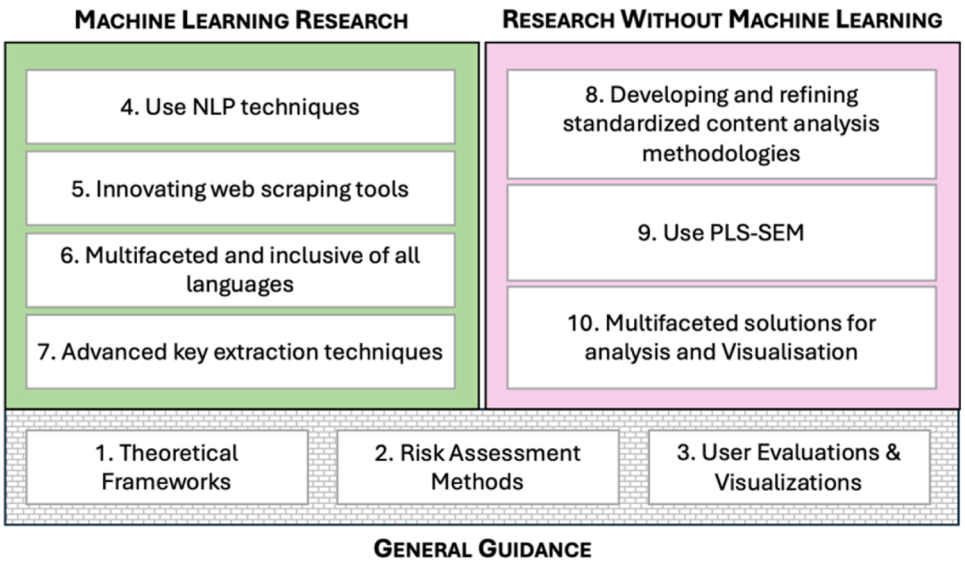

Fig. 2. Recommendation overview.

privacy risk assessment methods excel in systematically identifying potential privacy threats and vulnerabilities within policies. By utilizing these methods, researchers can meticulously analyze privacy policies, shedding light on areas where a user's data may be at risk. This precise identification is crucial for policymakers and organizations, as it enables them to proactively address potential privacy concerns, reducing the likelihood of data breaches or privacy violations. These risk assessment methods also enhance contextual understanding as they go beyond a superficial policy analysis. This latter aspect is crucial and clearly a significant challenge based on our review findings. Different industries and sectors have unique privacy requirements, and these methods facilitate the adaptation of evaluation criteria to suit specific contexts. As such, they enable researchers to assess the appropriateness of privacy measures within a given domain, resulting in more relevant and meaningful findings. Additionally, incorporating a DPIA as part of the privacy risk assessment process ensures a thorough examination of how data processing activities might impact the privacy of individuals. The DPIA evaluates not only the risks but also the necessity and proportionality of data processing activities, providing a clear framework for identifying and mitigating risks.

It is important to note that not all privacy risks are created equally and that their impact may vary significantly. The same applies to the selection of an appropriate method of assessment. First, researchers should clearly define the scope and objectives of their privacy assessment. This involves identifying the specific privacy risks they aim to evaluate, the context of the privacy policies (such as industry sector, geographical region, and type of data involved), and the regulatory requirements they need to comply with. Second, researchers should utilize comparative frameworks (or even decision-making tools) that evaluate various privacy risk assessment methods based on key criteria such as the applicability to diverse types of data, ease of use, comprehensiveness, and the level of detail they provide.

Either way, executing a formal privacy risk assessment enables researchers to prioritize privacy risks which in turn enhances an evaluation. Prioritizing in this manner is vital as it assists policymakers in allocating resources effectively. The same applies to compliance checking, regardless of the analysis technique used. Privacy risk assessment methods, including DPIAs, provide a structured framework for verifying whether privacy policies align with existing laws and regulations (e.g., GDPR). Such compliance checking frequently requires involving various stakeholders – an aspect many risk assessment methods are well equipped for. Many privacy risk assessment methods require the involvement of stakeholders, including users and experts, which ties in with our previous recommendation. Notably, the data-driven nature of privacy risk assessments, when combined with DPIAs, enables researchers to empirically prove the effectiveness of policies. Researchers can use the insights derived from these assessments to recommend policy improvements based on actual user data rather than relying solely on theoretical assumptions or subjective opinions. This is particularly important where emerging technologies such as IoT and blockchain are used, as a DPIA helps ensure that these innovations are deployed in ways that respect and protect user privacy.

Recommendation 3: To ensure the practicality, user-friendliness of visual representations, and comprehensiveness of privacy policy research, *it is imperative to improve certain methodological aspects. These include* ***incorporating user evaluations*** *and the prioritization of demographic (and geographic)* ***inclusivity when analyzing and collecting data for policy research.*** This cannot be overstated. Researchers should meticulously design and execute user studies to assess the effectiveness of visual representations in conveying intricate privacy information to users. These evaluations should include user comprehension, engagement, satisfaction, and trust in privacy design artifacts. Researchers can develop more user-centric privacy design solutions by systematically gathering user feedback and iteratively refining visual representations based on their input.

Moreover, conducting comparative studies evaluating the usability and effectiveness of different privacy design approaches is essential to identify best practices and areas that necessitate improvement. Datasets should also exhibit diversity by including policies from various sources, encompassing mobile apps and websites. Researchers should establish well-defined data collection protocols and annotation guidelines to mitigate dataset quality, transparency, and consistency issues. In longitudinal analysis, emphasis should be placed on capturing policy alterations over time, employing methods like version tracking or timestamping. Furthermore, researchers should consider exploring alternative data collection approaches, such as crowdsourcing, while remaining vigilant regarding potential annotation accuracy and participant engagement challenges. Additionally, research efforts should encompass a more comprehensive array of industries, demographics, and geographical regions.

4.2. Privacy policy research using machine learning

Given the technical nature of machine learning, we have tried to be as specific as possible.

Recommendation 4: Addressing the inherent complexity stemming

from ambiguity in privacy policies demands a concerted effort by advancing NLP techniques. *Researchers should therefore* ***prioritize the development of NLP models*** *adept at disambiguating and providing more precise interpretations of privacy policy text.* This necessitates the exploration of contextual embeddings such as BERT and extensive training on diverse privacy policy datasets (beyond OPP-115 and APP-350) to enhance the resolution of ambiguity. Additionally, attention should be given to incorporating techniques like coreference resolution and syntactic parsing to enhance the disambiguation process further. A critical aspect involves creating benchmark datasets tailored for evaluating and improving NLP model performance in disentangling ambiguity within privacy policies. Tangentially, researchers could make use of Named Entity Recognition (NER) to identify and categorize named entities within unstructured textual data, such as a natural language policy document. It is particularly adept at automating the entire process enabling organisations to perform compliance assessments, risk analyses, and benchmarking. Future research may wish to consider developing datasets to study cross-border data transfers. Here, NER could identify countries and the data transfer mechanisms used, such as the Standard Contractual Clauses (SCC), data protection adequacy decisions, and Binding Corporate Rules (BCR).

Recommendation 5: *To address the complexity associated with harvesting privacy policies, particularly in the context of longitudinal analysis, researchers should* ***focus on the development of innovative web scraping tools****.* These tools should be designed to extract structured and unstructured privacy policies while adapting to dynamic changes in website layouts and structures. Strategies involving web page monitoring and differential analysis should be explored to identify alterations in policies across various iterations. Furthermore, ensuring that harvested policies adhere to a standardized format or ontology is paramount to facilitate future analyses and comparisons. Legal experts should be included to maintain the quality and comprehensiveness of the harvested policies. Moreover, researchers should investigate the feasibility of harnessing blockchain or distributed ledger technology to establish a transparent and immutable record of policy alterations over time.

Recommendation 6: *Efforts to overcome language and semantic challenges in privacy policy analysis should be multifaceted and inclusive of non-English languages. As such, researchers* ***need to invest in refining machine translation systems capable of accurately translating privacy policies****, capturing the subtleties and cultural nuances inherent in the original language.* Leveraging multilingual NLP models trained on diverse datasets can significantly contribute to this endeavor. To address semantic challenges, researchers should develop methods to classify privacy policy text at the most suitable level of granularity, considering variations in language use. This includes creating models capable of discerning contextually relevant units (Grasso et al., 2024), such as paragraphs, sentences, or clauses, and integrating techniques like topic modeling and sentiment analysis to achieve a deeper understanding of the underlying semantics. Additional understanding could be gained by implementing state-of-the-art machine translation technologies. researchers should develop language-specific privacy ontologies. These ontologies are essential for standardizing the analysis process across different linguistic domains, providing a structured framework that can accommodate the unique characteristics of each language.

Researchers should also explore adapting pre-trained models, such as multilingual BERT, to enhance the accuracy of non-English policy analysis. These pre-trained models, designed to understand multiple languages, can be fine-tuned to better handle the specific nuances of privacy policies across various linguistic contexts. Additionally, developing bespoke language models for specific languages, particularly those that are less represented in existing datasets, can further improve translation and analysis accuracy. To assist in this regard, we advocate that researchers involve linguists who could further refine the translation processes whilst preserving the original meaning and cultural context. Similarly, legal experts can provide insights into the legal nuances that might differ across districts. Such an interdisciplinary approach would help develop a robust and reliable privacy policy analysis tool capable of working for a global audience. Moreover, researchers should consider the socio-cultural aspects that influence privacy perceptions in different regions. Understanding these cultural differences can aid in developing more nuanced models that respect and reflect the privacy concerns of diverse populations. This approach not only improves the accuracy of policy analysis but also enhances the relevance and sensitivity of privacy policies in various cultural contexts. In addition to technical advancements, fostering international collaborations and sharing resources across institutions can accelerate the development of comprehensive privacy policy analysis tools. By pooling expertise and data, researchers can create more sophisticated models that benefit from a wide array of linguistic and cultural inputs.

Recommendation 7: Keyword extraction remains a pivotal facet of privacy policy analysis, and its efficacy can be augmented by considering contextual information. *When extracting keywords, researchers should* ***explore advanced techniques considering context****, including the surrounding text, grammatical structure, and semantic relationships.* In addition to supervised methods, applying unsupervised learning techniques, such as word embeddings and topic modeling, should be explored to identify pertinent terms without explicit supervision. For dynamic analysis, where privacy policies evolve, researchers should investigate methods to detect emerging keywords and trends, ensuring that policy analyses remain current and reflect changing practices.

4.3. Privacy policy research without machine learning

Recommendation 8: The text highlights the prevalence of subjectivity in content analysis. *Researchers should* ***embark on developing and refining standardized content analysis methodologies*** *that mitigate subjectivity.* This entails investing in advanced NLP techniques, such as sentiment analysis and named entity recognition, to categorize and evaluate content objectively. Additionally, the utilization of machine learning algorithms for automated content analysis can contribute to consistent and reproducible results. Researchers should consider creating comprehensive guidelines for evaluators, including inter-rater reliability checks, to minimize interpretational variations. Collaboration with experts in NLP and content analysis is imperative to ensure the objectivity and reliability of the analytical process.

Recommendation 9: *We* ***recommend using PLS-SEM*** *as a suitable method for evaluating composite models and design artifacts that permeated the non-ML research reviewed* (Hair et al., 2019; Lowry and Gaskin, 2014). Several studies within our review developed a design artifact, but none conducted a formal user evaluation using PLS-SEM. We argue its suitability because PLS-SEM is particularly adept at:

- Situations where the relationships between variables (i.e., elements within a design artifact) are complex, non-linear, or not normally distributed. PLS-SEM can effectively capture and model these complex interactions when evaluating composite models, especially those involving intricate relationships among multiple constructs.
- Estimating models that contain reflective and formative constructs. In other words, PLS-SEM can estimate a model that includes constructs that are reflective (latent variables inferred from observed indicators) and formative (indicators defining the latent construct). This flexibility is crucial when dealing with composite models, as they frequently contain both constructs.
- Situations where measurement error is a consequence of the imperfect nature of artifact design, which translates to imperfectly measured variables or "noisy data."
- Estimating complex models that involve intricate paths with multiple mediation or moderation effects. This ability sets it apart from other methods of estimating structural equation models, such as (CB-SEM).
- Making use of data that is not normally distributed. Data is often non-normal in real-world scenarios, making a non-parametric technique, such as PLS-SEM, an appropriate choice.

- Integrating multi-method approaches where researchers have used various data collection methods (e.g., surveys, interviews, observations) in composite model evaluation.

Recommendation 10: *Visual data representation and* ***technical analysis*** *challenges* ***necessitate multifaceted solutions.*** These could entail the following:

- For visual representation, researchers should engage in user-centric design practices. Conducting user testing and feedback sessions on using icons and visual representations can guide the creation of more effective and user-friendly visual elements within privacy policies. Collaboration with user experience (UX) designers and graphic artists is essential to ensure the effectiveness of visual communication.
- To address scalability issues in visual data presentation, researchers should explore innovative data visualization techniques tailored to privacy policy content. Interactive visualizations and infographics can offer more accessible and scalable means of conveying complex information. Collaborating with data visualization experts can create dynamic and scalable visualizations that cater to a broad audience.
- Given the complexity of legal language in privacy policies, interdisciplinary collaboration with legal experts is indispensable. Legal professionals can provide crucial insights into the nuances of privacy policy content, ensuring that technical analysis accurately represents legal intricacies.
- Researchers should seek alternative data sources to circumvent technical constraints imposed by platforms like the Google Play Store. Innovative web scraping methods should be developed, adhering to platforms' terms of service and limitations on data crawling. Such procedures should ensure comprehensive access to data for analysis, mitigating potential technical bottlenecks and expanding the scope of research.

5. Limitations and future research

Although a rigorous study selection and analysis process was followed, this scoping review is not without limitations. The papers examined in this study were not heterogeneous; various methods and analytical approaches were used to produce their findings. This makes direct comparisons between papers challenging and limits the extent to which parallels can be drawn between publications. As in any scoping review, there is potential for researcher bias in selection or coding. Having said this, the use of independent coding (by two researchers) has reduced this potential within this context. This study did not incorporate grey literature. These sources may offer further insights given the growing interest in privacy and privacy policies. Finally, while this study offers recommendations for privacy policies and research, these have not been validated or tested.

Future research within this field should aim to address several key gaps and challenges identified in our review. First, there is a pressing need to develop standardized content analysis methodologies that minimize subjectivity. This includes leveraging advanced natural language processing techniques, such as sentiment analysis and named entity recognition, to provide objective categorizations and evaluations of privacy policy content. Additionally, researchers should explore the use of PLS-SEM to evaluate composite models and design artifacts more formally, as this method is particularly suited for complex, non-linear relationships, and measurement errors inherent in privacy policy research. Expanding the scope of data collection is also critical.

Future studies should include a broader range of industries, demographics, and geographic regions to ensure that findings are more generalizable and applicable to various contexts. Collaboration with industry partners, governmental agencies, and non-profit organizations can facilitate access to diverse datasets. Moreover, the development of comprehensive taxonomies or ontologies to classify privacy policies across different domains can help standardize analyses and comparisons. Addressing the visual and technical challenges in privacy policy analysis requires a multifaceted approach. User-centric design practices, including iterative user testing and feedback, can improve the effectiveness and accessibility of visual representations of privacy policies. Researchers should collaborate with UX designers, graphic artists, and legal experts to ensure that visual and technical analyses accurately reflect legal intricacies and user needs. Additionally, innovative data visualization techniques, such as interactive visualizations and infographics, can make complex information more accessible and scalable. The integration of advanced web scraping tools with sophisticated algorithms can enhance the extraction and analysis of privacy policies, particularly for longitudinal studies.

Lastly, future research should prioritize the inclusion of privacy-centric theoretical frameworks to provide a structured vocabulary and clear definitions, reducing ambiguity and ensuring the use of consistent terminology. Incorporating established privacy risk assessment methods can systematically identify potential privacy threats and vulnerabilities, enabling more precise and actionable recommendations for policymakers and organizations. By addressing these areas, future research can significantly advance the field of privacy policy analysis, making it more robust, comprehensive, and relevant to real-world applications.

6. Conclusion

The objective of this scoping review was to investigate privacy policy analysis more broadly by developing a holistic research agenda. To develop said research agenda, we set out to understand which analytic techniques are used when performing privacy policy analysis. Additionally, we wanted to understand which challenges these techniques embody when used to perform privacy policy analysis. Our findings emphasize the importance of grounding privacy policy analyses in established theoretical frameworks, fostering a comprehensive understanding of privacy concepts, and reducing ambiguity in research. Additionally, the integration of privacy risk assessment methods is underscored as a valuable approach for systematically identifying and addressing potential privacy threats within policies.

Furthermore, the review highlights the significance of user evaluations and user-centric design in the creation of effective visual representations of privacy policies, addressing both practicality and user-friendliness. For machine learning-based privacy policy analysis, the review recommends the use of advanced NLP techniques for disambiguating policy text. For research not utilizing machine learning, the scoping review advocates standardized content analysis methodologies to mitigate subjectivity. Furthermore, it suggests the utilization of PLS-SEM for evaluating composite models and design artifacts. Based on the above (amongst others) a set of research recommendations are discussed as part of our research agenda. Researchers and practitioners in the field can leverage these recommendations to enhance their analytical processes, resulting in more informed and actionable insights for privacy policy development and assessment.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Karl van der Schyff: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Project administration. **Suzanne Prior:** Formal analysis, Investigation, Writing – review & editing. **Karen Renaud:** Investigation, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cose.2024.104065.

References

Aberkane, A.J., Broucke, S.V., Poels, G., 2022. Investigating organizational factors associated with GDPR noncompliance using privacy policies: a machine learning approach. In: Proceedings of the 4th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA). IEEE, pp. 107–113. https://ieeexplore.ieee.org/document/10063341.

Acquisti, A., Brandimarte, L., Loewenstein, G., 2020. Secrets and likes: the drive for privacy and the difficulty of achieving it in the digital age. J. Consum. Psychol. 30 (4), 736–758. https://doi.org/10.1002/jcpy.1191.

Adhikari, A., Das, S., Dewri, R., 2022. Privacy policy analysis with sentence classification. In: Proceedings of the 19th Annual International Conference on Privacy, Security & Trust (PST). IEEE, pp. 1–10. https://ieeexplore.ieee.org/document/9851977.

Ahmad, W.U., Chi, J., Tian, Y., Chang, K.W., 2020. PolicyQA: a reading comprehension dataset for privacy policies. arXiv preprint arXiv:2010.02557. doi:10.48550/arXiv.2010.02557.

Ahmad, J., Li, F., Luo, B., Atluri, V., Di Pietro, R., Jensen, C.D., Meng, W., 2022. IoTPrivComp: a measurement study of privacy compliance in IoT apps. In: Computer Security. Lecture Notes in Computer Science, 13555. Springer, Cham. https://doi.org/10.1007/978-3-031-17146-8_29.

Akanfe, O., Valecha, R., Rao, H.R., 2020. Assessing country-level privacy risk for digital payment systems. Comput. Secur. 99, 102065 https://doi.org/10.1016/J.COSE.2020.102065.

Alabduljabbar, A., Abusnaina, A., Meteriz-Yildiran, Ü., Mohaisen, D., 2021a. Automated privacy policy annotation with information highlighting made practical using deep representations. In: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, pp. 2378–2380. https://doi.org/10.1145/3460120.3485335.

Alabduljabbar, A., Abusnaina, A., Meteriz-Yildiran, Ü., Mohaisen, D., 2021b. TLDR: deep learning-based automated privacy policy annotation with key policy highlights. In: Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society, pp. 103–118. https://doi.org/10.1145/3463676.3485608.

Alshamsan, A.R., Chaudhry, S.A., 2022. Machine learning algorithms for privacy policy classification: a comparative study. In: Proceedings of the 2nd International Conference on Software Engineering and Artificial Intelligence (SEAI), pp. 214–219. https://doi.org/10.1109/SEAI55746.2022.9832027.

Al Rahat, T.A., Long, M., Tian, Y, 2022. Is your policy compliant? A deep learning-based empirical study of privacy policies' compliance with GDPR. In: Proceedings of the 21st Workshop on Privacy in the Electronic Society, pp. 89–102. https://doi.org/10.1145/3559613.3563195.

Amos, R., Acar, G., Lucherini, E., Kshirsagar, M., Narayanan, A., Mayer, J., 2021. Privacy policies over time: curation and analysis of a million-document dataset. In: Proceedings of the 2021 Web Conference, pp. 2165–2176. https://doi.org/10.1145/3442381.3450048.

Andow, B., Mahmud, S., Wang, W., Whitaker, J., Enck, W., Reaves, B., Singh, K., 2019. PolicyLint: investigating internal privacy policy contradictions on google play. In: Proceedings of the 28th USENIX Security Symposium, pp. 585–602.

Arksey, H., O'Malley, L, 2005. Scoping studies: towards a methodological framework. Int. J. Soc. Res. Methodol. 8 (1), 19–32. https://doi.org/10.1080/1364557032000119616.

Asif, M., Javed, Y., Hussain, M., 2021. Automated analysis of pakistani websites' compliance with GDPR and Pakistan data protection act. In: Proceedings of the 2021 International Conference on Frontiers of Information Technology (FIT), pp. 234–239. https://doi.org/10.1109/FIT53504.2021.00051.

Audich, D.A., Dara, R., Nonnecke, B., Gal-Oz, N., Lewis, P., 2018. Privacy policy annotation for semi-automated analysis: a cost-effective approach. In: Trust Management XII. IFIPTM 2018. IFIP Advances in Information and Communication Technology, 528. Springer, Cham. https://doi.org/10.1007/978-3-319-95276-5_3.

Audich, D.A., Dara, R., Nonnecke, B., 2021. Improving readability of online privacy policies through DOOP: a domain ontology for online privacy. Digital 1 (4), 198–215. https://doi.org/10.3390/digital1040015.

Bachiri, M., Idri, A., Fernández-Alemán, J.L., Toval, A., 2018. Evaluating the privacy policies of mobile personal health records for pregnancy monitoring. J. Med. Syst. 42 (8), 144. https://doi.org/10.1007/s10916-018-1002-x.

Bareh, C.K., 2022. Privacy policy analysis for compliance and readability of library vendors in India. Ser. Libr. 83 (2), 148–165. https://doi.org/10.1080/0361526X.2022.2143467.

Bateni, N., Kaur, J., Dara, R., Song, F., 2022. Content analysis of privacy policies before and after GDPR. In: Proceedings of the 19th Annual International Conference on Privacy, Security & Trust (PST), pp. 1–9. https://doi.org/10.1109/PST55820.2022.9851983.

Becher, S.I., Benoliel, U., Mathis, K., Tor, A., 2021. Law in books and law in action: the readability of privacy policies and the GDPR. In: Consumer Law and Economics. Economic Analysis of Law in European Legal Scholarship, 9. Springer, Cham. https://doi.org/10.1007/978-3-030-49028-7_9.

Bhatia, J., Breaux, T.D., Schaub, F., 2016. Mining privacy goals from privacy policies using hybridized task recomposition. ACM Trans. Softw. Eng. Methodol. 25 (3), 1–24. https://doi.org/10.1145/2907942.

Bhattacharjee, K., Chen, M., Dasgupta, A., 2020. Privacy-preserving data visualization: reflections on the state of the art and research opportunities. Comput. Graph. Forum 39 (3), 675–692. https://doi.org/10.1111/cgf.14032.

Boliek, B.E., 2021. Upgrading unconscionability: a common law ally for a digital world. Md. Law Rev. 81. https://heinonline.org/HOL/Page?handle=hein.journals/mllr81&id=54&div=&collection=.

Bolton, T., Dargahi, T., Belguith, S., Maple, C., 2023. PrivExtractor: toward redressing the imbalance of understanding between virtual assistant users and vendors. ACM Trans. Priv. Secur. 26 (3), 1–29. https://doi.org/10.1145/3588770.

Bookert, N., Bondurant, W., Anwar, M., 2022. Data practices of internet of medical things: a look from privacy policy perspectives. Smart Health 26, 100342. https://doi.org/10.1016/j.smhl.2022.100342.

Bowers, J., Reaves, B., Sherman, I.N., Traynor, P., Butler, K., 2017. Regulators, mount Up! analysis of privacy policies for mobile money services. In: Proceedings of 13th Symposium on Usable Privacy and Security (SOUPS 2017), pp. 97–114.

Bowers, J., Sherman, I.N., Butler, K.R.B., Traynor, P., 2019. Characterizing security and privacy practices in emerging digital credit applications. In: Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks, pp. 94–107. https://doi.org/10.1145/3317549.3319723.

Bracamonte, V., Hidano, S., Tesfay, W., Kiyomoto, S., 2019. Evaluating privacy policy summarization: an experimental study among Japanese users. In: Proceedings of the 5th International Conference on Information Systems Security and Privacy, pp. 370–377. https://doi.org/10.5220/0007378403700377.

Bui, D., Tang, B., Shin, K.G., 2022. Do opt-outs really opt me out?. In: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, pp. 425–439. https://doi.org/10.1145/3548606.3560574.

Bui, D., Yao, Y., Shin, K.G., Choi, J.M., Shin, J., 2021. Consistency analysis of data-usage purposes in mobile apps. In: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, pp. 2824–2843. https://doi.org/10.1145/3460120.3484536.

Buitelaar, J.C., 2017. Post-mortem privacy and informational self-determination. Ethics Inf. Technol. 19 (2), 129–142. https://doi.org/10.1007/s10676-017-9421-9.

Cadogan, R.A., 2011. An imbalance of power: the readability of internet privacy policies. J. Bus. Econ. Res. JBER 2 (3). https://doi.org/10.19030/jber.v2i3.2864.

Chang, K.C., Zaeem, R.N., Barber, K.S., Susilo, W., Deng, R.H., Guo, F., Li, Y., Intan, R., 2020. A framework for estimating privacy risk scores of mobile apps. In: Information Security. ISC 2020. Lecture Notes in Computer Science, 12472. Springer, Cham. https://doi.org/10.1007/978-3-030-62974-8_13.

Chang, C., Li, H., Zhang, Y., Du, S., Cao, H., Zhu, H., Biagioni, E., Zheng, Y., Cheng, S., 2019. Automated and personalized privacy policy extraction under GDPR consideration. In: Wireless Algorithms, Systems, and Applications. Lecture Notes in Computer Science, 11604. Springer, Cham. https://doi.org/10.1007/978-3-030-23597-0_4.

Chaw, C.Y., Chua, H.N., Kim, H., Kim, K.J., 2021. A framework system using word mover's distance text similarity algorithm for assessing privacy policy compliance. In: IT Convergence and Security. Lecture Notes in Electrical Engineering, 782. Springer, Singapore. https://doi.org/10.1007/978-981-16-4118-3_8.

Costante, E., Sun, Y., Petković, M., Den Hartog, J., 2012. A machine learning solution to assess privacy policy completeness: (short paper). In: Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society, pp. 91–96. https://doi.org/10.1145/2381966.2381979.

Cottrill, C., Thakuriah, P., 2011. Protecting location privacy: policy evaluation. Transp. Res. Rec. 2215 (1), 67–74.

Cooke, M., Buckley, N., 2008. Web 2.0, social networks and the future of market research. Int. J. Mark. Res. 50 (2), 267–292. https://doi.org/10.1177/147078530805000208.

Cui, H., Trimananda, R., Markopoulou, A., Jordan, S., 2023. PoliGraph: automated privacy policy analysis using knowledge graphs. In: Proceedings of the 32nd USENIX Security Symposium (USENIX Security 23), pp. 1037–1054.

Del Alamo, J.M., Guaman, D., Balmori, B., Diez, A., 2021. Privacy assessment in android apps: a systematic mapping study. Electronics 10 (16), 1999. https://doi.org/10.3390/electronics10161999 (Basel).

Del Alamo, J.M., Guaman, D.S., García, B., Diez, A., 2022. A systematic mapping study on automated analysis of privacy policies. Computing 104 (9), 2053–2076. https://doi.org/10.1007/s00607-022-01076-3.

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., Colautti, C., 2006. Privacy calculus model in E-commerce – a study of Italy and the United States. Eur. J. Inf. Syst. 15 (4), 389–402. https://doi.org/10.1057/palgrave.ejis.3000590.

Dombetzki, L., Kecht, C., Kratsch, W., Rau, D., 2020. Amaryllis: a user-centric information system for automated privacy policy analysis. In: Proceedings of the 28th European Conference on Information Systems, pp. 1–18. https://aisel.aisnet.org/ecis2020_rp/83.

Drozd, O., Kirrane, S., Hölbl, M., Rannenberg, K., Welzer, T., 2020. Privacy CURE: consent comprehension made easy. In: ICT Systems Security and Privacy Protection. SEC 2020. IFIP Advances in Information and Communication Technology, 580. Springer, Cham. https://doi.org/10.1007/978-3-030-58201-2_9.

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

Ebert, N., Ackermann, K.A., Scheppler, B., 2021. Bolder is better: raising user awareness through salient and concise privacy notices. In: Proceedings of the 2021 Conference on Human Factors in Computing Systems, 12, pp. 1–12. https://doi.org/10.1145/3411764.3445516.

Efroni, Z., Metzger, J., Mischau, L., Schirmbeck, M., 2019. Privacy icons: a risk-based approach to visualisation of data processing. Eur. Data Prot. Law Rev. 5, 352.

Farooq, E., Nawaz UI Ghani, M.A., Naseer, Z., Iqbal, S, 2020. Privacy policies' readability analysis of contemporary free healthcare apps. In: Proceedings of the 14th International Conference on Open Source Systems and Technologies (ICOSST), pp. 1–7. https://doi.org/10.1109/ICOSST51357.2020.9332991.

Fox, G., Lynn, T., Rosati, P., 2022. Enhancing consumer perceptions of privacy and trust: a GDPR label perspective. Inf. Technol. People 35 (8), 181–204.

Fox, G., Tonge, C., Lynn, T., Mooney, J., 2018. Communicating compliance: developing a GDPR privacy label. In: Proceedings of the 24th Americas Conference on Information Systems, pp. 1–5.

Gerl, A., Meier, B., 2019. Privacy in the future of integrated health care services-are privacy languages the key?. In: Proceedings of the International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 312–317. https://doi.org/10.1109/WIMOB.2019.8923532.

Ghazinour, K., Albalawi, T., 2016. A usability study on the privacy policy visualization model. In: Proceedings of the 14th International Conference on Dependable, Autonomic and Secure Computing, pp. 578–585. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2016.109.

Grasso, B., La Gatta, V., Moscato, V., Sperlì, G., 2024. KERMIT: knowledge-empowered model in harmful meme detection. Inf. Fusion 106, 102269. https://doi.org/10.1016/j.inffus.2024.102269.

Guntamukkala, N., Dara, R., Grewal, G., 2015. A machine-learning based approach for measuring the completeness of online privacy policies. In: Proceedings of the 14th International Conference on Machine Learning and Applications (ICMLA). IEEE, pp. 289–294. https://doi.org/10.1109/ICMLA.2015.143.

Guo, W., Rodolitz, J., Birrell, E., 2020. Poli-See: an interactive tool for visualizing privacy policies. In: Proceedings of the 19th Workshop on Privacy in the Electronic Society, pp. 57–71. https://doi.org/10.1145/3411497.3420221.

Guamán, D.S., Rodriguez, D., del Alamo, J.M., Such, J., 2023. Automated GDPR compliance assessment for cross-border personal data transfers in android applications. Comput. Secur. 130, 103262.

Hair, J.F., Risher, J.J., Sarstedt, M., Ringle, C.M., 2019. When to use and how to report the results of PLS-SEM. Eur. Bus. Rev. 31 (1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203.

Hamdani, R.El, Mustapha, M., Amariles, D.R., Troussel, A., Meeùs, S., Krasnashchok, K, 2021. A combined rule-based and machine learning approach for automated GDPR compliance checking. In: Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law, pp. 40–49. https://doi.org/10.1145/3462757.3466081.

Harkous, H., Fawaz, K., Lebret, R., Schaub, F., Shin, K., Aberer, K., 2018. Polisis: automated analysis and presentation of privacy policies using deep learning. In: Proceedings of the 27th USENIX Security Symposium, pp. 531–548.

Hashmi, S.S., Waheed, N., Tangari, G., Ikram, M., Smith, S., Hara, T., Yamaguchi, H., 2022. Longitudinal compliance analysis of android applications with privacy policies. In: Mobile and Ubiquitous Systems: Computing, Networking and Services. MobiQuitous 2021. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 419. Springer, Cham. https://doi.org/10.1007/978-3-030-94822-1_16.

Hatamian, M., Momen, N., Fritsch, L., Rannenberg, K., Naldi, M., Italiano, G., Rannenberg, K., Medina, M., Bourka, A., 2019. A multilateral privacy impact analysis method for android apps. In: Privacy Technologies and Policy. APF 2019. Lecture Notes in Computer Science, 11498. Springer, Cham. https://doi.org/10.1007/978-3-030-21752-5_7.

Hatamian, M., Wairimu, S., Momenluca, N., et al., 2021. A privacy and security analysis of early-deployed COVID-19 contact tracing android apps. Empir. Softw. Eng. 26, 36. https://doi.org/10.1007/s10664-020-09934-4.

Javed, Y., Al Qahtani, E., Shehab, M., 2021. Privacy policy analysis of banks and mobile money services in the middle east. Future Internet 13 (1), 10. https://doi.org/10.3390/fi13010010.

Jilka, S., Simblett, S., Odoi, C.M., van Bilsen, J., Wieczorek, A., Erturk, S., Wykes, T., 2021. Terms and conditions apply: critical issues for readability and jargon in mental health depression apps. Internet Interv. 25, 100433.

John, N.A., 2013. Sharing and Web 2.0: the emergence of a keyword. New Media Soc. 15 (2), 167–182. https://doi.org/10.1177/1461444812450684.

Kandil, S.A., van den Akker, M., van Baarsen, K., Jansen, S., van Vulpen, P., Wnuk, K., Brinkkemper, S., 2018. Benchmarking privacy policies in the mobile application ecosystem. In: Software Business. ICSOB 2018. Lecture Notes in Business Information Processing, 336. Springer, Cham. https://doi.org/10.1007/978-3-030-04840-2_4.

Kaplan, S., Bulmer, D., Gosselin, A., Ghanavati, S., 2021. Lattice-based contextual integrity analysis of social network privacy policies. In: Proceedings of the 29th International Requirements Engineering Conference Workshop (REW), pp. 394–399. https://doi.org/10.1109/REW53955.2021.00070.

Kelley, P.G., Bresee, J., Cranor, L.F., Reeder, R.W., 2009. A "Nutrition label" for privacy. In: Proceedings of the 5th Symposium on Usable Privacy and Security, pp. 1–12. https://doi.org/10.1145/1572532.1572538.

Kelley, P.G., Cesca, L., Bresee, J., Cranor, L.F., 2010. Standardizing privacy notices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1573–1582. https://doi.org/10.1145/1753326.1753561.

Kelley, P.G., Cranor, L.F., Sadeh, N., 2013. Privacy as part of the app decision-making process. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3393–3402. https://doi.org/10.1145/2470654.2466466.

Khandelwal, R., Nayak, A., Chung, P., & Kassem, F., 2023. The overview of privacy labels and their compatibility with privacy policies. arXiv:2303.08213. doi:10.48550/arXiv.2303.08213.

Kitchenham, B., Brereton, P., Li, Zhi, Budgen, D., Burn, A., 2011. Repeatability of systematic literature reviews. In: Proceedings of the 15th Annual Conference on Evaluation & Assessment in Software Engineering (EASE 2011), pp. 46–55. https://doi.org/10.1049/ic.2011.0006.

Kotal, A., Joshi, A., Joshi, K.P., 2021. The effect of text ambiguity on creating policy knowledge graphs. In: Proceedings of the 2021 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking. IEEE, pp. 1491–1500. https://doi.org/10.1109/ISPA-BDCloud-SocialCom-SustainCom52081.2021.00201.

Kretschmer, M., Pennekamp, J., Wehrle, K., 2021. Cookie banners and privacy policies: measuring the impact of the GDPR on the web. ACM Trans. Web TWEB (4), 15. https://doi.org/10.1145/3466722.

Lau, J., Zimmerman, B., Schaub, F., 2018. Alexa, are you listening?. In: Proceedings of the ACM on Human-Computer Interaction, pp. 1–31. https://doi.org/10.1145/3274371.

Laufer, R.S., Wolfe, M., 1977. Privacy as a concept and a social issue: a multidimensional developmental theory. J. Soc. Issues 33 (3), 22–42. https://doi.org/10.1111/j.1540-4560.1977.tb01880.x.

Leicht, J., Gerl, A., & Heisel, M. (2021). Technical report on the extension of the layered privacy language. https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en.

Leicht, J., Heisel, M., 2019. A survey on privacy policy languages: expressiveness concerning data protection regulations. In: Proceedings of the 12th CMI Conference on Cybersecurity and Privacy (CMI), pp. 1–6. https://doi.org/10.1109/CMI48017.2019.8962144.

Levac, D., Colquhoun, H., O'Brien, K.K, 2010. Scoping studies: advancing the methodology. Implement. Sci. 5 (1), 1–9. https://doi.org/10.1186/1748-5908-5-69/TABLES/3.

Liao, S., Wilson, C., Cheng, L., Hu, H., Deng, H., 2020. Measuring the effectiveness of privacy policies for voice assistant applications. In: Proceedings of the Annual Computer Security Applications Conference, pp. 856–869. https://doi.org/10.1145/3427228.3427250.

Libert, T., 2018. An automated approach to auditing disclosure of third-party data collection in website privacy policies. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18, pp. 207–216. https://doi.org/10.1145/3178876.3186087.

Liepin, R., Contissa, G., Drazewski, K., Lagioia, F., Lippi, M., Micklitz, H., Palka, P., Sartor, G., Torroni, P., 2019. GDPR privacy policies in CLAUDETTE: challenges of omission, context and multilingualism. In: Proceedings of the 3rd Workshop on Automated Semantic Analysis of Information in Legal Texts, pp. 1–7.

Lin, X., Liu, H., Li, Z., Xiong, G., Gou, G., 2022. Privacy protection of China's top websites: a multi-layer privacy measurement via network behaviours and privacy policies. Comput. Secur. 114, 102606 https://doi.org/10.1016/j.cose.2022.102606.

Liu, B., Ding, M., Shaham, S., Rahayu, W., Farokhi, F., Lin, Z., 2022. When machine learning meets privacy. ACM Comput. Surv. 54 (2), 1–36. https://doi.org/10.1145/3436755.

Liu, F., Ramanath, R., Sadeh, N., Smith, N., 2014. A step towards usable privacy policy: automatic alignment of privacy statements. In: Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers, pp. 884–894.

Liu, K., Xu, G., Zhang, X., Xu, G., Zhao, Z., 2022. Evaluating the privacy policy of android apps: a privacy policy compliance study for popular apps in China and Europe. Sci. Program. 2022, 1–15. https://doi.org/10.1155/2022/2508690.

Liu, S., Zhang, F., Zhao, B., Guo, R., Chen, T., Zhang, M., 2023. APPCorp: a corpus for android privacy policy document structure analysis. Front. Comput. Sci. 17 (3), 173320 https://doi.org/10.1007/s11704-022-1627-2.

Lowry, P.B., Gaskin, J., 2014. Partial Least Squares (PLS) Structural Equation Modeling (SEM) for building and testing behavioral causal theory: when to choose it and how to use it. IEEE Trans. Prof. Commun. 57 (2), 123–146. https://doi.org/10.1109/TPC.2014.2312452.

Manandhar, S., Kafle, K., Andow, B., Singh, K., Nadkarni, A., 2022. Smart home privacy policies demystified: a study of availability, content, and coverage. In: Proceedings of the 31st USENIX Security Symposium, pp. 3521–3538.

McDonald, A.M., Reeder, R.W., Kelley, P.G., Cranor, L.F., Goldberg, I., Atallah, M.J., 2009. A comparative study of online privacy policies and formats. In: Privacy Enhancing Technologies. PETS 2009. Lecture Notes in Computer Science, 5672. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-03168-7_3.

McRobb, S., Rogerson, S., 2004. Are they really listening? Inf. Technol. People 17 (4), 442–461. https://doi.org/10.1108/09593840410570285.

Meier, Y., Schäwel, J., Krämer, N.C., 2020. The shorter the better? Effects of privacy policy length on online privacy decision-making. Media Commun. 8 (2), 291–301. https://doi.org/10.17645/MAC.V8I2.2846.

Mohammadi, N.G., Pampus, J., Heisel, M., 2019. Pattern-based incorporation of privacy preferences into privacy policies. In: Proceedings of the 24th European Conference

on Pattern Languages of Programs, pp. 1–12. https://doi.org/10.1145/3361149.3361154.
Mousavi Nejad, N., Jabat, P., Nedelchev, R., Scerri, S., Graux, D., Hölbl, M., Rannenberg, K., Welzer, T., 2020. Establishing a strong baseline for privacy policy classification. In: ICT Systems Security and Privacy Protection. SEC 2020. IFIP Advances in Information and Communication Technology, 580. Springer, Cham. https://doi.org/10.1007/978-3-030-58201-2_25.
Mousavi Nejad, N., Scerri, S., Lehmann, J., Faron Zucker, C., Ghidini, C., Napoli, A., Toussaint, Y., 2018. KnIGHT: mapping privacy policies to GDPR. In: Knowledge Engineering and Knowledge Management. EKAW 2018. Lecture Notes in Computer Science, 11313. Springer, Cham. https://doi.org/10.1007/978-3-030-03667-6_17.
Munn, Z., Peters, M.D.J., Stern, C., Tufanaru, C., McArthur, A., Aromataris, E., 2018. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. BMC Med. Res. Methodol. 18 (1), 143. https://doi.org/10.1186/s12874-018-0611-x.
Narksenee, M., Sripanidkulchai, K., 2019. Can we trust privacy policy: privacy policy classification using machine learning. In: Proceedings of the 2nd International Conference of Intelligent Robotic and Control Engineering (IRCE), pp. 133–137. https://doi.org/10.1109/IRCE.2019.00034.
Neal, D., Gaber, S., Joddrell, P., Brorsson, A., Dijkstra, K., Dröes, R.M., 2023. Read and accepted? Scoping the cognitive accessibility of privacy policies of health apps and websites in three European Countries. Digit. Health 9. https://doi.org/10.1177/20552076231152162/SUPPL_FILE/SJ-XLSX-1-DHJ-10.1177_20552076231152162.XLSX.
Nissenbaum, H., 2019. Contextual integrity up and down the data food chain. Theor. Inq. Law 20 (1), 221–256. https://doi.org/10.1515/til-2019-0008.
Nokhbeh Zaeem, R., Anya, S., Issa, A., Nimergood, J., Rogers, I., Shah, V., Barber, K.S., 2020. PrivacyCheck v2: a tool that recaps privacy policies for you. In: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, pp. 3441–3444. https://doi.org/10.1145/3340531.3417469.
Nokhbeh Zaeem, R., Barber, K.S., 2021. A large publicly available corpus of website privacy policies based on DMOZ. In: Proceedings of the 11th ACM Conference on Data and Application Security and Privacy, pp. 143–148. https://doi.org/10.1145/3422337.3447827.
Novikova, E., Doynikova, E., Kotenko, I., Katsikas, S., et al., 2020. P2Onto: making privacy policies transparent. In: Computer Security. Lecture Notes in Computer Science, 12501. Springer, Cham. https://doi.org/10.1007/978-3-030-64330-0_15.
Obar, J.A., Oeldorf-Hirsch, A., 2020. The biggest lie on the internet: ignoring the privacy policies and terms of service policies of social networking services. Inf. Commun. Soc. 23 (1), 128–147. https://doi.org/10.1080/1369118X.2018.1486870.
Oltramari, A., Piraviperumal, D., Schaub, F., Wilson, S., Cherivirala, S., Norton, T.B., Russell, N.C., Story, P., Reidenberg, J., Sadeh, N., 2018. PrivOnto: a semantic framework for the analysis of privacy policies. Semant. Web 9 (2), 185–203. https://doi.org/10.3233/SW-170283.
Paspatis, I., Tsohou, A., Kokolakis, S., 2020. AppAware: a policy visualization model for mobile applications. Inf. Comput. Secur. 28 (1), 116–132. https://doi.org/10.1108/ICS-04-2019-0049.
Paul, N., Tesfay, W.B., Kipker, DK., Stelter, M., Pape, S., Janczewski, L., Kutyłowski, M., 2018. Assessing privacy policies of internet of things services. In: ICT Systems Security and Privacy Protection. SEC 2018. IFIP Advances in Information and Communication Technology, 529. Springer, Cham. https://doi.org/10.1007/978-3-319-99828-2_12.
Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D., Soares, C.B., 2015. Guidance for conducting systematic scoping reviews. Int. J. Evid. Based Healthc. 13 (3), 141–146. https://doi.org/10.1097/XEB.0000000000000050.
Pham, M.T., Rajić, A., Greig, J.D., Sargeant, J.M., Papadopoulos, A., Mcewen, S.A., 2014. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res. Synth. Methods 5 (4), 371–385. https://doi.org/10.1002/JRSM.1123.
Qiu, W., Lie, D., 2020. Deep active learning with crowdsourcing data for privacy policy classification. arXiv:2008.02954. doi:10.48550/arXiv.2008.02954.
Reeder, R.W., Kelley, P.G., McDonald, A.M., Cranor, L.F., 2008. A user study of the expandable grid applied to P3P privacy policy visualization. In: Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society, pp. 45–54. https://doi.org/10.1145/1456403.1456413.
Reeder, R.W., Kelley, P.G., McDonald, A.M., Cranor, L.F., 2009. A user study of the expandable grid applied to P3P privacy policy visualization. In: Proceedings of the 5th Symposium on Usable Privacy and Security. https://doi.org/10.1145/1572532.1572582.
Sanghavi, P., Ghamsani, R., Parekh, R., Mota, R., Dongre, D., 2022. Simplifying privacy agreements using machine reading comprehension and open domain. In: Proceedings of the 6th International Conference on Computing, Communication, Control and Automation, ICCUBEA 2022. https://doi.org/10.1109/ICCUBEA54992.2022.10010822.
Sarne, D., Schler, J., Singer, A., Sela, A., Bar Siman Tov, I., 2019. Unsupervised topic extraction from privacy policies. In: Proceedings of the 2019 World Wide Web Conference, pp. 563–568. https://doi.org/10.1145/3308560.3317585.
Shayegh, P., Jain, V., Rabinia, A., Ghanavati, S., 2019. Automated approach to improve IoT privacy policies. arXiv:1910.04133. doi:10.48550/arXiv.1910.04133.
Shvartzshnaider, Y., Apthorpe, N., Feamster, N., Nissenbaum, H., 2019. Going against the (Appropriate) flow: a contextual integrity approach to privacy policy analysis. In: Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, pp. 162–170. https://doi.org/10.1609/hcomp.v7i1.5266.
Solove, D.J., 2021. The myth of the privacy paradox. George Wash. Law Rev. 89. https://heinonline.org/HOL/Page?handle=hein.journals/gwlr89&id=15&div=&collection=.
Soumelidou, A., Tsohou, A., 2020. Effects of privacy policy visualization on users' information privacy awareness level. Inf. Technol. People 33 (2), 502–534. https://doi.org/10.1108/ITP-08-2017-0241.
Story, P., Zimmeck, S., Ravichander, A., Smullen, D., Wang, Z., Reidenberg, J., Sadeh, N., 2019. Natural language processing for mobile app privacy compliance. In: Proceedings of the AAAI Spring Symposium on Privacy-enhancing Artificial Intelligence and Language Technologies, 2, p. 4.
Subahi, A., Theodorakopoulos, G., Cagáňová, D., Horňáková, N., 2023. Automated Approach to Analyze IoT Privacy Policies. Industry 4.0 Challenges in Smart Cities. EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-92968-8_12.
Sunyaev, A., Dehling, T., Taylor, P.L., Mandl, K.D., 2015. Availability and quality of mobile health app privacy policies. J. Am. Med. Inform. Assoc. 22 (e1), e28–e33. https://doi.org/10.1136/amiajnl-2013-002605.
Tesfay, W.B., Hofmann, P., Nakamura, T., Kiyomoto, S., Serna, J., 2018. I read but don't agree. In: Proceedings of the 2018 Web Conference, pp. 163–166. https://doi.org/10.1145/3184558.3186969.
Thotawaththa, T.A.I., Gamage, Y.T., Gamlath, D., Chee, W., Meedeniya, D., 2021. Automated categorization of privacy policies based on user perspective. In: Proceedings of the 10th International Conference on Information and Automation for Sustainability (ICIAfS), pp. 54–59. https://doi.org/10.1109/ICIAfS52090.2021.9606158.
Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D.J., Horsley, T., Weeks, L., Hempel, S., Akl, E.A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M.G., Garritty, C., Straus, S.E., 2018. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann. Intern. Med. 169 (7), 467–473. https://doi.org/10.7326/M18-0850.
Trimananda, R., Le, H., Cui, H., Ho, J.T., Shuba, A., Markopoulou, A., 2022. OVRseen: auditing network traffic and privacy policies in oculus VR. In: Proceedings of the 31st USENIX Security Symposium (USENIX Security 22, pp. 3789–3806.
Tucker, R., Tucker, C., Zheng, J., 2015. Privacy pal: improving permission safety awareness of third-party applications in online social networks. In: Proceedings of the 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems, pp. 1268–1273. https://doi.org/10.1109/HPCC-CSS-ICESS.2015.83.
Usableprivacy.org. (2017, May 27). Usableprivacy.org: the usable privacy project.
Vanezi, E., Zampa, G., Mettouris, C., Yeratziotis, A., Papadopoulos, G.A., Cherfi, S., Perini, A., Nurcan, S., 2021. CompLicy: evaluating the GDPR alignment of privacy policies - a study on web platforms. In: Research Challenges in Information Science. RCIS 2021. Lecture Notes in Business Information Processing, 415. Springer, Cham. https://doi.org/10.1007/978-3-030-75018-3_10.
Wagner, I., 2023. Privacy policies across the ages: content of privacy policies 1996–2021. ACM Trans. Priv. Secur. 26 (3), 1–32. https://doi.org/10.1145/3590152.
West, T., 2022. Children's privacy: an evaluation of EdTech privacy policies. In: Proceedings of the Conference on Information Systems Applied Research, pp. 1–12.
Wettlaufer, J., Simo, H., Friedewald, M., Önen, M., Lievens, E., Krenn, S., Fricker, S., 2020. Decision support for mobile app selection via automated privacy assessment. Privacy and Identity Management. Data for Better Living: AI and Privacy. Springer, pp. 292–307. https://doi.org/10.1007/978-3-030-42504-3_19.
Wilson, S., Schaub, F., Ramanath, R., Sadeh, N., Liu, F., Smith, N.A., Liu, F., 2016. Crowdsourcing annotations for websites' privacy policies. In: Proceedings of the 25th International Conference on World Wide Web, pp. 133–143. https://doi.org/10.1145/2872427.2883035.
Yang, L., Chen, X., Luo, Y., Lan, X., Chen, L., 2021. PurExt: automated extraction of the purpose-aware rule from the natural language privacy policy in IoT. Secur. Commun. Netw. 2021, 1–11. https://doi.org/10.1155/2021/5552501.
Yu, L., Luo, X., Liu, X., Zhang, T., 2016. Can we trust the privacy policies of android apps?. In: Proceedings of the 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 538–549. https://doi.org/10.1109/DSN.2016.55.
Yuan, H., Boakes, M., Ma, X., Cao, D., Li, S., Cabanillas, C., Pérez, F., 2023. Visualising personal data flows: insights from a case study of booking.com. In: Intelligent Information Systems. CAiSE 2023. Lecture Notes in Business Information Processing, 477. Springer, Cham. https://doi.org/10.1007/978-3-031-34674-3_7.
Zhao, W., Shahriar, H., Clincy, V., Bhuiyan, Z.A., 2020. Security and privacy analysis of Mhealth application: a case study. In: Proceedings of the 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), pp. 1882–1887. https://doi.org/10.1109/TrustCom50675.2020.00257.
Zhu, R., Srivastava, A., Sutanto, J., 2020. Privacy-deprived E-commerce: the efficacy of consumer privacy policies on China's E-commerce websites from a legal perspective. Inf. Technol. People 33 (6), 1601–1626. https://doi.org/10.1108/ITP-03-2019-0117/FULL/PDF.
Zimmeck, S., Bellovin, S., 2014. Privee: an architecture for automatically analyzing web privacy policies. In: Proceedings of the 23rd USENIX Security Symposium (USENIX Security 23, pp. 1–17.
Zimmeck, S., Story, P., Smullen, D., Ravichander, A., Wang, Z., Reidenberg, J., Cameron Russell, N., Sadeh, N., 2019. MAPS: scaling privacy compliance analysis to a million apps. In: Proceedings of the Privacy Enhancing Technologies, 2019, pp. 66–86. https://doi.org/10.2478/popets-2019-0037.

Karl van der Schyff: Karl van der Schyff holds a BSc, MSc, and PhD. His research interests focus on the behavioral and psychological dimensions of information security and digital forensics. He has authored and co-authored multiple peer-reviewed publications and has reviewed works for top-tier IS journals.

Suzanne Prior Dr Suzanne Prior is a lecturer in the Division of Cybersecurity at Abertay University. She has an undergraduate in Applied Computing and a PhD, both from the University of Dundee. Prior's current research interests surround usability challenges involved in cyber security, focusing on the suitability of cybersecurity mechanisms for children and how to best educate children on cybersecurity issues.

Karen Renaud Karen Renaud is a Scottish computing Scientist at the University of Strathclyde in Glasgow, working on all aspects of Human-Centred Security and Privacy. She is particularly interested in deploying behavioural science techniques to improve security behaviours, and in encouraging end-user privacy-preserving behaviours.