"It doesn't tell me anything about how my data is used": User Perceptions of Data Collection Purposes

Lin Kyi
Max Planck Institute for Security and Privacy
Bochum, Germany
lin.kyi@mpi-sp.org

Abraham Mhaidli
University of Michigan
Ann Arbor, United States
mhaidli@umich.edu

Cristiana Santos
Utrecht University
Utrecht, The Netherlands
c.teixeirasantos@uu.nl

Franziska Roesner
University of Washington
Seattle, United States
franzi@cs.washington.edu

Asia Biega
Max Planck Institute for Security and Privacy
Bochum, Germany
asia.biega@mpi-sp.org

ABSTRACT

Data collection purposes and their descriptions are presented on almost all privacy notices under the GDPR, yet there is a lack of research focusing on how effective they are at informing users about data practices. We fill this gap by investigating users' perceptions of data collection purposes and their descriptions, a crucial aspect of informed consent. We conducted 23 semi-structured interviews with European users to investigate user perceptions of six common purposes (*Strictly Necessary, Statistics and Analytics, Performance and Functionality, Marketing and Advertising, Personalized Advertising,* and *Personalized Content*) and identified elements of an effective purpose name and description.

We found that most purpose descriptions do not contain the information users wish to know, and that participants preferred some purpose names over others due to their perceived transparency or ease of understanding. Based on these findings, we suggest how the framing of purposes can be improved toward meaningful informed consent.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Security and privacy** → **Usability in security and privacy**; • **Applied computing** → **Law**.

KEYWORDS

GDPR, personal data, purposes, privacy, tracking, qualitative methods

ACM Reference Format:
Lin Kyi, Abraham Mhaidli, Cristiana Santos, Franziska Roesner, and Asia Biega. 2024. "It doesn't tell me anything about how my data is used": User Perceptions of Data Collection Purposes. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3613904.3642260

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642260

1 INTRODUCTION

With the enforcement of the European Union's General Data Protection Regulation (GDPR) [19] in 2018 and the ePrivacy Directive [23], privacy notices (also known as cookie banners or consent notices) have become the de facto standard for informing and collecting consent from EU and UK users. This has created a new industry of GDPR compliance tools, such as IAB Europe's Transparency and Consent Framework (TCF) and Consent Management Platforms (CMPs), established to help organizations manage their GDPR and ePD compliance through privacy notices [53]. Additionally, several governing authorities, such as national Data Protection Authorities (DPAs) have been formed and have set up guidelines for GDPR compliance.

Due to the ubiquity of consent dialogs, users in the EU and UK are now generally familiar with this process of consenting *to something*, but do they actually know *what* they are consenting to? The UK and EU GDPR mandates that consent be *informed* [19], yet many studies on privacy notices have shown that being informed is often simply *assumed* [53, 65].

At the core of informed consent lie the data collection purposes for which users are sharing their data for. Examples of such purposes presented in privacy notices may include: "Strictly Necessary", "Advertising", "Analytics", etc. as illustrated in Figure 1. Current regulatory guidelines for how data processing purposes should be described or named vary significantly [49, 55], therefore diversity exists between various websites, DPA guidelines, and CMP templates. Problems also exist within these purpose names and descriptions, such as: how these names do not always accurately map onto the technical services provided [9], or how, as we have observed in this paper, they can exploit cognitive biases. Some purpose names and descriptions are trickier for users to understand than others. Overall, whether purpose formulations are effective at informing users about what their data would be used for remains an understudied topic.

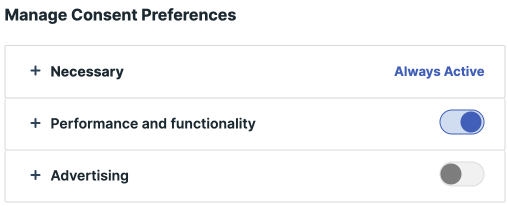

Figure 1: An example of a privacy notice with its listed data collection purposes.

Focusing on the data collection purposes and identifying the major issues with them are important for three reasons. First, although there is research showing that users are often the weakest point in security [32, 56], there will always be a small group of users who want to be informed about their privacy, and take their time to make privacy-conscious choices. Currently, privacy notices are riddled with deceptive designs [41, 49, 53, 65], therefore we believe that reframing these purposes will make online data collection practices more transparent and informative for these users. Second, privacy nutrition labels have shown that when privacy information is presented to users in a digestible manner, users often are more informed without feeling overwhelmed [37, 38]. Therefore, we lay the groundwork for future work on redesigning more user-centric consent systems. Third, reframing data collection purposes allows for more efficient management and repurposing of data [2, 7], when combined with appropriate technical and legal measures. Many organizations are not fully aware of what is happening with the data they collect [3], which is problematic from the perspective of user privacy.

There is a complex network of adtech vendors and third parties involved in the collection and processing of user data [66], but users tend to group third parties, service providers, and other data controllers as being the same entity [45]. Moreover, privacy notices are an important mechanism wherein information about purposes is disclosed to users. There is value in understanding how users perceive these purposes, and understanding how we can make them more user-friendly because users are requested to make decisions regarding these purposes every time they are online. We focused on HCI-specific contributions, but suggest that a transdisciplinary solution encompassing HCI, the law, and technical stakeholders be involved to drive meaningful change.

We thus argue that more attention needs to be paid to the data collection purposes and the text within the privacy notices to develop better formulations that accurately and effectively inform users about how their data is being processed. Hence, we investigated how users perceive the data collection purposes they might be consenting to, and identified elements of an effective purpose name and description to better improve the *informed* aspect of informed consent.

The research questions in this paper are:

(1) How do users evaluate common data collection purposes and their descriptions?
(2) How do users prefer data collection purposes be named and described so that they are more user-friendly?

To answer these research questions, we carried out interviews with European internet users ($n = 23$) to understand how they perceived common data collection purposes. We found that most purpose descriptions do not contain information participants want to know, including how long their data is retained for, and how to request their data be deleted. For purpose names, participants found some names to be more clear and understandable compared to others. Participants had varying opinions for the different purposes presented; some purposes, such as *Strictly Necessary* purposes being were more accepted than *Personalized Advertising* for sharing data with. *Statistics and Analytics* or *Performance and Functionality* purposes were not commonly understood properly. Based on our findings, combined with research insights from psychology, we propose guidelines to improve data collection purpose names and descriptions, which may likely improve how *informed* users are in the informed consent process.

2 BACKGROUND

We describe the legal and design foundations for our work. There are few legal specifications about what purposes data can be collected for, how they should be described, and how to best inform users before collecting their consent, which impacts the usability of privacy notices on a deeper level beyond the UI.

2.1 Legal Background

The GDPR applies to the processing of personal data [8] of EU and UK users and requires organizations to choose a legal basis to lawfully process personal data (Article 6(1)(a)). When the legal basis being applied is consent, the GDPR also defines the requirements for a valid consent. Article 5(1)(a) and Recital 60 of the GDPR also require disclosure of information which is triggered by the principles of lawfulness, fairness, and transparency. The ePrivacy Directive (ePD) provides supplementary rules to the GDPR in particular for the use of tracking technologies, such as cookies. Article 5(3) of the ePrivacy Directive requires websites to give clear and comprehensive information when requesting consent for non-necessary tracking purposes for the service requested by the use (such as targeting advertising, social networks, third-party analytics).

Some purposes are exempt from consent, such as *functional* or *essential* trackers (Recital 66 ePD). The only way to assess with certainty whether consent is required is to analyze the purpose of each tracker on a given website [15, 16, 25]. Cookie purposes allowing website owners to retain the preferences expressed by users, regarding a service, should be deemed essential or technically necessary. However, research has also found that sometimes purposes deemed essential are being used for non-essential purposes, such as advertising [9].

2.1.1 Scope of the GDPR in the UK. As our participant pool consists of UK residents in addition to EU residents, we address the scope of the GDPR in the UK. As a result of Brexit, the EU GDPR is not in effect within the UK, however, the provisions of the EU GDPR were incorporated directly into the UK law as the "UK GDPR"[64]. In practice, there is little change to the core data protection principles, rights and obligations, and organizations can operate as they did pre-Brexit [35]. Additionally, the UK DPA maintains that they work closely with the EU for data protection [35]. Thus, we use the term

"GDPR" throughout the paper to refer to both the UK and the EU GDPR.

2.1.2 Legal requirements for purpose formulation. Pursuant to the principle of purpose limitation (Article 5(1)(b) GDPR [15]), personal data can be collected for specified, explicit and legitimate purposes only. Santos et al. studied various legal documents and systematized the legal requirements for purposes, which require i) *explicitness* (availability, unambiguity, shared common understanding), ii) *specificity*, iii) *intelligible* (non technical terms, conciseness), iv) *clear and plain language* (straight forward statements, concreteness), and concrete requirements for consent: v) *freely given* and vi) *informed* consent [55].

Most important to our work is the *informed consent* requirement. Whenever tracking technologies are deployed on a user's device, the user must be given clear and comprehensive information, and the content information must comprise the purposes of processing and the means for expressing their consent, pursuant to Article 5(3) of the ePD. The need to present information on the processing operations is triggered by the principles of lawfulness, fairness, and transparency depicted in Article 5(1)(a) and the recitals of the GDPR. In particular, Recital 60 explains that "a data controller should provide a data subject with all information necessary to ensure fair and transparent processing, taking into account the specific circumstances (…)."

2.1.3 DPAs require clear purposes. Regulatory guidelines provide examples of purposes, yet there is no consensus on which formulation of purposes is preferred. As such, several national DPAs have different standards and guidelines for organizations subject to the GDPR. The Italian DPA confirms the absence of a standardized naming convention for cookies' purposes [36].

The UK DPA acknowledges that while providing information about cookies' purposes equates to transparency requirements, users may not always understand that information. The UK DPA encourages websites to make an effort to explain their activities in an understandable manner, but it does not impose strict requirements [33]. The Latvian DPA requires that the information provided not contain unduly legal or technical language [46].

The French DPA recommends formulating purposes in a descriptive and intuitive name so that users can be fully aware of the possibility of exercising a choice by purpose. It says that purposes should be formulated "in an intelligible way, in a suitable language and clear enough to allow users to understand precisely what they are consenting to." It also recommends that each purpose be highlighted in a short and highlighted title, accompanied by a brief description [12].

In addition, the EDPB Taskforce [22] acknowledges that some service providers classify "essential" or "strictly necessary" cookies and processing operations which would not be considered as "strictly necessary" within the meaning of Article 5(3) ePD or the ordinary meaning of "strictly necessary" or "essential" under the GDPR.

2.2 Related Work on Data Collection Purposes

In addition to guidelines from the various DPAs regarding which purposes to use, data controllers need to consider whether to leave purposes general, therefore giving users fewer choices for control, or more specific, therefore giving users more choice.

Utz et al.'s analysis of privacy notice interfaces found that 45.5% of banners used generic purposes, such as "improving user experience", 38.6% used specific purposes, such as "ad delivery", and 16.9% did not even mention their purposes [65]. Korff et al. found that when participants were presented with more privacy setting choices, they were less happy and more likely to regret their choice [39].

When presented with more specific purpose choices, Habib et al. found that users are more likely to accept only *Strictly Necessary* cookies or make more granular consent choices if the UI made it easy to do so [28]. They studied user comprehension of four purposes categories developed by The UK International Chamber of Commerce: i) *Strictly Necessary*, ii) *Performance*, iii) *Functionality*, and iv) *Targeting/Advertising*, and found that the purpose categories of *Performance* and *Functionality* were the most misunderstood by users [28].

Previous work has shown that the design of a privacy notice impacts how users make consent choices [28, 49, 53]. Service providers, consent management platforms (CMPs), and third party vendors, commonly use deceptive practices to collect users' consent, such as by making it difficult to reject consent, and not properly informing users [41, 49, 53, 65]. As such, many users find privacy notices to be annoying, and do not pay much attention to them [43].

Bouma-Sims et al. found that few users actually read purpose definitions, though no significant difference in comprehension was noted when definitions were provided [11]. Kyi et al. found that users tended to be most accepting of sharing data for *Strictly Necessary, Security and Debugging,* and *Fraud and Law Enforcement* legitimate interest purposes, but least accepting of sharing data for *Personalized Ads* and *Sharing Data with Third Parties* legitimate interest purposes [45].

Not only must organizations consider the number of purposes presented in privacy notices, but they should also consider users' perceptions of these data collection purposes. The data collection purposes offered, and their applications are often a multi-stakeholder situation, involving many different actors, such as IAB Europe and CMPs in addition to the service provider [31, 45].

2.3 Deceptive Design Beyond User Interfaces

While there has been a plethora of research looking at deceptive design choices in the user interfaces of privacy notices [45, 49, 53, 65], less attention has been paid on deceptive practices outside of the UI [45, 55]. Of the work that has looked into non-UI deceptive designs, studies have shown that deceptive practices go well beyond the UI, such as deceptive linguistic practices [45, 48, 55].

Previous work by Santos et al. has found that 89% of privacy notices violate the GDPR; 61% were too vague in describing purposes (thereby violating the *purpose specificity* principle), and 30% framed their data practices in positive language (violating the *freely given and informed* requirements for consent) [55]. Kyi et al. found that linguistic deceptive designs were exploited in the use of legitimate interests, such as by providing placebic and/or positive explanations to users about their data collection practices, and being vague about legal terms [45].

Habib et al. and Ma et al. found that loss aversion text, which is where privacy notice text might point out the negative outcomes of not accepting all cookies, was influential in making users believe they had to accept all cookies [28, 48]. This practice should therefore be avoided as a practice by data controllers [28]. Berens et al. also confirmed this finding, showing in their study that the phrasing for accepting or rejecting cookies can influence users' behaviours [6].

Related work in this space suggests that descriptions and linguistic elements, such as positive or negative framing and being transparent about practices, can impact users' consent choices and perceptions. We extend upon this work by looking at the linguistic elements (i.e., purpose names and descriptions) of privacy notices, and the challenges and opportunities for consent that they may present.

3 METHODS

We conducted semi-structured interviews with 23 English-speaking participants over 18 years old from the UK and Ireland. We recruited participants using Prolific[1], a website for recruiting online participants for research studies. Since many languages are spoken within the countries scoped by the GDPR, and different languages might present different nuances in the text of a privacy notice, we recruited from the UK and Ireland to increase the chances that participants were exposed to English-language privacy notices, and thus more familiar with the data collection purposes we presented to them.

Interviews took approximately one hour to complete, after which participants were compensated €23 for their time. Our Institutional Review Board declared that this study was exempt from ethical review, however we had to collect consent from participants for the audio recordings to be GDPR-compliant. Interviews were conducted during June 2023. After the interviews were conducted, we used an automatic tool to transcribe our interview audio, and two researchers annotated the interviews. As our study spans across multiple disciplines, we had a multidisciplinary team, consisting of those from computer science, psychology, HCI, and legal backgrounds.

3.1 Selecting Data Collection Purposes

As purposes are flexible, and there is no standardized list of purposes and their descriptions, we had to search through various DPA guidelines, Consent Management Platforms (CMPs), and the ePrivacy Directive to collect purposes and descriptions for the second and third sections of our interview. Regulators propose various ways to formulate purposes. Some DPAs, such as those discussed in Section 2.1, provide more concrete guidelines about how to name and describe purposes, which has helped to guide our study.

We initially collected a set of 201 purposes and descriptions from 6 CMPs (OneTrust, Quantcast, TrustArc, Cookiebot, LiveRamp, and Crownpeak) and 39 companies from our own online browsing and looking at previous literature studying data collection purposes [28, 31, 55]. We ultimately decided on a smaller set of six purposes for conducting the interviews because we noticed that most purposes we collected were CMP- and DPA-based purposes. Therefore, focusing on CMP- and DPA-based purposes provided a wider coverage of purposes and descriptions that are used in practice [31].

The six data collection purposes we decided on are:

(1) *Strictly Necessary / Essential / Required*;
(2) *Performance / Functionality*;
(3) *Statistics / Analytics*;
(4) *Advertising*;
(5) *Personalized Advertising*; and
(6) *Personalized Content*

The different descriptions corresponding to each purpose are included in Section 2 of our Supplementary Materials.

We decided on these six purposes because we wanted purposes that are both widely used and broad enough that they covered a variety of different uses. As a point of comparison, some of these purposes have been studied in other papers [28, 45], but we also added other purposes that have not been studied yet to gain new insights. While *Advertising* and *Personalized Advertising* seem similar, we wanted to see whether participants could differentiate between the two, and how they felt about many purposes with different names being related to advertising.

We pilot tested our interviews with three participants; all regularly used the internet in English, one participant was from a computational background, one from a non-technical privacy background, and one from a non-technical background for variety. The pilot tests helped us reformulate our interview questions, and indicated that showing participants six data collection purposes, each having between three to five different descriptions, was within participants' attention limits.

3.2 Interview Procedure

During the interviews, we first introduced our work and the research team, then asked participants to fill out a consent form and a demographics form. Thereafter, we gave participants definitions of what a "data collection purpose" meant, and provided screenshots of data collection purposes in privacy notices to provide more context. We then proceeded with the interview questions, which we summarize below (see Section 1 of our Supplementary Materials for the full interview protocol). All interviews were conducted by the first author on Zoom, with the option for participants to turn on their video.

Our semi-structured interview consisted of three sections. In the first section, participants were asked about what they expect organizations to disclose in privacy notices (Q1.1), whether they think it is necessary to share their data with organizations (Q1.2), how they feel about privacy notices telling them of services they will miss out on if they declined cookies (Q1.3), and how well-informed they feel about online data practices (Q1.4).

In the second section, participants were presented with the names of six data collection purposes without any definitions (*Strictly Necessary / Essential / Required, Performance and Functionality, Statistics and Analytics, Advertising, Personalized Advertising,* and *Personalized Content*). For each purpose, we asked what they think happens with their data under this purpose (Q2.1), whether they would feel comfortable sharing data for that purpose (Q2.2), and what they think would happen if they denied consent for that purpose (Q2.3). We repeated this set of questions for each purpose.

[1] prolific.co

In the third section, participants were given a link to a document that showed each of the six data collection purposes presented in section two of the interview, and presented with various descriptions from different sources to read (see Section 2 of our Supplementary Materials for these descriptions). After participants read the description for one purpose, they were asked how they would describe that purpose in their own words (Q3.1), how similar they felt the descriptions were (Q3.2), whether there was a description they preferred the most (Q3.3), how well-informed they felt about how that purpose uses their data (Q3.4), what could be improved in the descriptions (Q3.5), how clear the purpose name was in describing what it does (Q3.6), and if the name was not clear, whether they had suggestions for a better or improved name (Q3.7). We repeated the procedure of having participants read the descriptions and answer these questions for each of our six purposes.

3.2.1 Participants. We recruited participants who were over 18, spoke and used the internet primarily in English, and lived in the UK or Ireland to ensure exposure to English privacy notices. Despite Brexit, the UK GDPR remains largely similar to the EU GDPR, and UK residents are still subjected to responding to privacy notices [34]. Therefore, UK residents should be just as familiar with privacy notices as EU residents, and including UK residents would give us access to a larger pool of potential research participants. We stopped recruiting at 23 participants because we reached saturation by this point, meaning we stopped hearing new topics being brought up at this point [26, 57]. Most interviews reach saturation between 9 and 17 interviews, and for studies with a general population, such as ours, saturation is reached at approximately 16 interviews [30].

Our participants were roughly split between men (48%) and women (52%), all primarily used the internet in English, and all participants except for one had been living in the UK or Ireland for over four years. We did not collect information about participants' educational backgrounds, but did aim for a representative sample across gender and age, therefore we had a wide variety of ages represented in our sample; 17% were between the ages of 18 to 24, another 17% were between 25 to 34 years old, 26% between 35 to 44 years old, 23% between 45 to 54 years old, and 17% were over 55 years old. Most of our participants came from non-computational backgrounds.

3.3 Data Analysis

Codebook. Our qualitative interview data was analyzed by two authors of this paper, starting with an inductive, open-ended approach to data analysis, then a deductive approach with a codebook that was revised during the annotation process [62]. This method was preferred since the interview investigated different data collection purposes, allowing for us to connect codes to specific purposes we studied. See Section 3 of our Supplementary Materials for our codebook.

Annotations. Since the first annotator conducted all the interviews, they initially open-coded three interviews to form an initial codebook. Afterwards, the first and second annotator met to annotate another three interviews together, discuss, and adjust the codebook as needed. The annotators then coded the same set of another six interviews separately, meeting after each interview to compare codes and discuss. After they reached an interrater reliability (IRR) of 80% ($kappa$ = 0.79), which indicates substantial agreement [50], the annotators split the rest of the interviews and annotated them separately. For increased validity, the annotators individually re-annotated the initial interviews where IRR was not yet reached.

4 RESULTS

In this section we describe our qualitative results. As our interviews were not directly measuring quantities, we have used terminology present in previous qualitative HCI studies to give estimates of quantities [18, 28, 45].

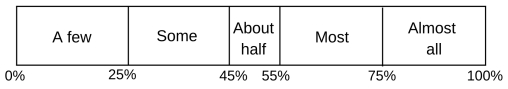

Figure 2: Terminology used to represent the frequency of themes in our qualitative results. This graphic was taken from [45].

4.1 Perceptions of Data Collection Practices

During the interview, participants provided comments about their thoughts regarding data collection practices in general.

Participants did not feel informed about data collection practices. None of the participants felt well-informed of online data practices, even if the law prescribes the variety of information an organization needs has to disclose to users to ensure fair and transparent processing (Articles 5(1)(a), 14, Recital 39 GDPR).

Almost all participants said they did not know i) what data is being collected, ii) how it is collected, iii) why their data is collected, iv) to whom this data is being sent to, nor the v) sensitivity of the data being collected. Additionally, many participants said they do not trust the information websites share about how they process user data, believing that all purposes were being used covertly for advertising in some way. Accordingly, mandated informational and transparency requirements are not efficient for meaningful choices.

Participants saw privacy notices negatively. Very few participants read privacy notices regularly; almost all participants felt privacy notices were annoying and just usually do what it takes to get rid of them quickly, such as clicking "Accept all" or accepting only necessary cookies by default. This echoes findings from previous research which found that users tend to not interact with privacy notices very thoroughly [11, 28, 53].

Participants believed that sharing their data is unnecessary, or a trade-off. When asked how necessary participants thought it was to share their data with organizations, we received a mixed response. Some participants believed it was necessary to share as much data as organizations are collecting in exchange for using these free services.

However, many others believed it was not necessary to share as much data as organizations are (perceived to be) collecting because these are often over-collecting data. As one participant stated, *"I don't think it's necessary, but it's just becoming more the norm nowadays. Most people just don't really know that actually their data is being used or stored at all"* (P2).

Some participants said that while sharing data is not necessary, they feel resigned to share it because organizations have more power in the end. They felt they have little control over the data they can share with organizations. Even if users wanted to decline sharing data, most believed organizations would still find a way to collect their information, echoing a finding from Kulyk et al. [42].

4.2 Perceptions of Data Collection Purposes

Herein we discuss participants' general perceptions of the six data collection purposes we presented: i) *Strictly Necessary / Required / Essential*, ii) *Performance / Functionality*, iii) *Statistics / Analytics*, iv) *Advertising*, v) *Personalized Advertising*, and vi) *Personalized Content*.

Most participants believed nothing would happen if they declined tracking. When asked what they think would happen if they declined tracking for a purpose, most participants believed it would make no noticeable impact on their current online experience. This finding is somewhat contrary to previous work which showed that only a few participants believed nothing would happen by declining tracking [42]. This difference may be due to recent case-law compelling companies to present more balanced options in consent notices, so that users reject tracking more easily in consent notices and notice not much happens when tracking is declined [13]. As P23 described, *"I have declined them sometimes, but I've never noticed any difference between declining them or accepting them."* Other participants' quotes supplement this finding: *"I don't think anything would happen really, except maybe you'd end up with less spam"* (P2) and *"I think nothing would probably happen, except that the provider would probably be less than happy because you'd be less of a useful customer to them"* (P9).

Some others believed they might get fewer services relating to the particular purpose they declined, such as less targeted advertising if they declined *Personalized Advertising*, or would not be able to access the site depending on the purpose(s) they declined. A few participants said they lacked the technical knowledge to assess how site functionalities would be impacted if they declined cookies, therefore felt the need to accept all tracking.

Participants either felt threatened or inquisitive about missing services because they rejected tracking. When asked how they felt about privacy notices that tell them they are missing out on certain services by rejecting tracking, we had mixed responses from participants. Some participants found this information important, and wanted to know about the services that they would be missing out on if they declined purposes.

Conversely, some believed they did not need to know what they were missing out on and felt these messages were threatening, as if organizations were forcing them to share their data. As stated by P4, *"It feels like a manipulation, if I'm honest. It's kind of a way of them trying to get you to consent to things."*

Most participants believed *Strictly Necessary* purposes were mandatory, but some participants had doubts. Most participants understood *Strictly Necessary* purposes as being mandatory, and that they have to accept it to access a website. However, we also had some participants who doubted whether this purpose was actually necessary for site access. This is a correct assumption according to Article 5(3) ePD. The assumption of necessity requires a technical analysis to assess whether any tracker is indeed necessary for a website to work [9].

Additionally, some participants believed this purpose was a vague, "catch all" kind of purpose. As explained by P7: *"I don't think it's necessary, but it does seem that most of the defaults are to allow everything. On a lot of sites, it does seem where it's only strictly necessary cookies (being shown to users). So it seems like they they're just relying on people just to click 'accept' without really looking at any of the other details."* Participants sometimes described that this purpose provided advertising or provided a better site experience for users, conflating it with other purposes such as *Performance and Functionality* purposes.

Participants linked the majority of purposes to advertising. Most participants correctly thought *Performance and Functionality* purpose provided them with more effective services and remembered their choices. Yet, some participants also believed this purpose provides them with targeted ads, as explained by P22, *"I would have guessed this similar to Statistic and Analytics, which I would like to call 'advertising.' I don't know what I would expect to be different."*

Statistics and Analytics purpose were commonly believed to be used to analyze data from users using the site, and to analyze website statistics for future improvements. In line with perceptions from other purposes we presented, some participants believed *Statistics and Analytics* was yet another purpose being used for sending marketing and advertising materials to users.

Some participants conflated *Personalized Content* with *Personalized Advertising* purposes. However, most participants believed *Personalized Content* was meant for creating user profiles and showing personalized content based on these profiles. Regarding the creation of user profiles, many participants said they wanted more information from organizations about how they were creating their profiles. As explained by P19, *"It just tells me that they will give me content based on what I like, it doesn't tell me anything about how my data is used."*

***Advertising* was conflated with *Personalized Advertising*.** In the collection of purposes presented, we found that certain DPAs (French and Spanish DPAs), and CMPs (OneTrust, TrustArc, and LiveRamp) differentiated between *Advertising* and *Personalized Advertising* purposes. *Personalized Advertising* and *Advertising* purposes were always confused with each other by participants, wherein they believed that both of these purposes were meant to serve personalized ads. In reality, *Advertising* is used for delivering generic ads to users. All participants were correct in their interpretation of *Personalized Advertising* purposes, believing it is used to serve users personalized ads.

Participants were not comfortable with sharing data for *Advertising* purposes. As for user comfort with sharing data for the six purposes we presented to participants, it differed by purpose. Most participants were not happy sharing data for *Personalized Advertising* and *Advertising* purposes. Instead, participants said that they were more comfortable sharing their data for *Strictly Necessary*, *Performance and Functionality*, and *Statistics and Analytics* purposes. When they found out what *Personalized Content* purposes did, and realized it was not the same as *Personalized Advertising*, most participants said they were comfortable sharing their data for this purpose due to the perceived convenience of this purpose, echoing previous research findings [40, 45].

4.3 Perceptions of Purpose Names

When asked how clear the name of a given purpose was, participants had some mixed responses regarding certain purposes. When it came to suggesting concrete names to make them more comprehensible, participants were generally much better at knowing why a name was unclear rather than suggesting a new and improved name.

***Personalized Advertising* is a clear name.** This name was the clearest purpose name to participants compared to the other purposes we presented. All of our participants felt the name was expressive of what the purpose does, and participants' perceptions of the purpose matched what purpose descriptions said. On the flipside, almost all participants conflated *Advertising* for *Personalized Advertising* purposes, believing them to provide the same functions.

Some purposes make use of conjunction and synonymy. Sometimes, purposes are called by several names by different companies, DPAs, or CMPs, such as *"Strictly Necessary / Required / Essential"*. Additionally, purposes may also use conjunctions, which is when several functionalities are combined into one purpose, such as *"Statistics and Analytics"* or *"Performance and Functionality."* Notably, regulators also use synonyms to name purposes, such as using "Technical / Required / Functional cookies" as per the Latvian DPA's guidelines [47]. Conjunctions, such as "Personalised ads and content, ad and content measurement, audience insights and product development" are used by CMPs like LiveRamp.

Where conjunction names were presented, participants often had a preference for one of these names, finding it clearer and/or more transparent than the other(s), as explained below.

Strictly Necessary / Required / Essential. Some participants wondered if these purposes were actually "strictly necessary" for a website to work. In the words of P6, *"I think it (the name, Strictly Necessary / Required / Essential) needs to be improved because 'Necessary for who?' is my question. It might be necessary for the person that wants your information, but not to me as a person using it. I would like a little bit of clarification on who it's necessary for because from the title it looks as though it's just necessary for the organisation and not for me."*

Statistics and Analytics. Some participants found *Analytics* to be more transparent and descriptive compared to *Statistics*. As described by P18, *"I would go with 'Analytics' rather than 'Statistics'. For me, it's just knowing how they sort of use the data. I don't need to know the figures and stuff."*

Performance and Functionality. Most participants found *Performance* to be a better name. To many participants, *Functionality* implies that this purpose is important, or even necessary, for the website to function, whereas *Performance* aligns better with the descriptions of this purpose, which relates to UX improvements and site experience. As P20 explains, *"'Functional' to me just means that it makes it (the service) work, whereas 'Performance' indicates that it will try to be as best as it can, like how good is the performance."*

4.4 Perceptions of Purpose Descriptions

We presented multiple descriptions from various sources (the ePrivacy Directive, DPAs and CMPs) for each purpose. Based on participant responses after viewing these descriptions, we gained insight into what can be done to improve purpose descriptions.

Participants preferred simple descriptions with less technical jargon. Most participants preferred simpler descriptions, but with more relevant and concise information about how their data is processed, as described in Section 4.1. Further, they preferred when descriptions avoided technical and legal jargon, such as terms like "persistent cookies" and "SSO". As P15 says, *"It talks about single sign on (SSO), and I feel like, like... as someone who doesn't really know this term, I wouldn't really want to see the 'SSO' part. I wouldn't want to make it seem as though it's harder to read than it actually is."* Consent choices in current interfaces often employ language that is too technical or confusing for users [59], which contradicts legal requirements that mandate consent collection to be understood by an "average member of the intended audience" [15, 20].

Participants suggested that privacy notices should be more visually appealing to capture attention. To improve readability and capture user attention towards data collection purposes and descriptions, some participants suggested that privacy notices describe purposes in point-form, and made better use of colour and icons to make them more appealing, rather than the current paragraphs of text describing purposes.

Participants want to be further informed about data retention and data rights. In addition to being better informed of what data is being collected for and who it is sent to, participants also wanted to be informed of the time their data is retained for once they accept tracking for a purpose, what happens to their data once they end the session, and the sensitivity of the data that is being collected about them.

Participants also mentioned wanting to know more about how they can request organizations to have their data deleted, what happens to their data if they reject cookies, and the services that are still provided once they reject cookies. As described by P3, *"(I want to know) how I can go and remove it (my data) if I've already consented to it. It would be nice if these parts of data collection were disclosed here (the privacy notice) as well."*

Participants wanted more reassurance in the description of purposes. Most participants preferred when purpose descriptions provided them with reassurance that their data would be kept safe, such as when descriptions mention that user data would be kept anonymous, or not be used for profiling purposes. Given the lack of transparency in how user data is handled, participants preferred it when companies gave them reassurance about keeping their data private. As P15 puts it, *"It mentioned they don't directly store personal information, which is quite nice to know."*

The need for reassurance was also observed by some participants who said they would like organizations to better describe the technical elements behind their data collection, such as by providing hyperlinks to find more information, or providing brief definitions for technical terms. Despite these suggestions, participants also said they would be unlikely to read or find out more about technical terms, but they want to feel reassured to know that there is the option to find out more information if they chose to. The EDPB guidelines [21] match user perceptions suggesting the use of technical definitions and examples.

Descriptions for the same purpose can be perceived as dissimilar. When presented with several different descriptions for the same purpose, participants often felt descriptions were different due to three reasons. First, participants indicated a disparity in the depth

of information given; some descriptions were more informative than others, therefore were perceived to be more transparent. Second, a difference resided in the examples purpose descriptions gave. Participants noted that the examples of services provided would vary, making them confused about what the purpose actually did. Last, participants sometimes commented that these different descriptions were sometimes giving them conflicting information. This was especially prominent in *Statistics and Analytics*, and also *Performance and Functionality* purposes, a finding Habib et al. also found [28]. As demonstrated by P3, "I feel like they have slightly different definitions (for *Performance and Functionality)*. It kind of takes me back to *Statistics and Analytics* a little bit where we were talking about the performance over there as well."

Different purpose descriptions can be perceived as describing the same functions. Participants sometimes mentioned that, based on descriptions for different purposes, some purposes seemed similar because of the services mentioned and perceived functions of that purpose. For example, some participants pointed out that two descriptions for *Statistics and Analytics* mentioned advertising, even though it is not an advertising purpose. Similarly, participants commonly confused *Statistics and Analytics*, *Performance and Functionality*, and *Strictly Necessary* purposes with each other, believing their descriptions sometimes overlapped.

5 DISCUSSION

In this paper we investigated two research questions: i) how users evaluate commonly used data collection purposes and their descriptions, and ii) how users prefer data collection purposes be named and described so that they are more effective.

We expand on the knowledge of informed consent by focusing on ways that users can be better informed about online data collection purposes and their descriptions. Based on our findings, in addition to insights from research in psychology, we present several recommendations to further improve the framing of data collection purposes towards informed consent.

5.1 Towards User-Informed Purpose Names

Based on previous research and our findings, we make the following recommendations for reframing data collection purpose names.

A middle ground is needed between overly broad and overly narrow purposes. We caution against using overly broad and overly narrow purposes, which Machuletz et al. also suggested in their work [49]. Websites and regulators need to find a middle ground among the various purpose names. Overly broad purposes, such as presenting only *Strictly Necessary* and *Non-essential* purposes were deemed too vague and broad for users to fully understand what is happening with their data, and gives users less choice about their data. This is used in the case of the Italian DPA that only distinguishes between two broad categories of "technical cookies", and "profiling cookies" [36]. Overly narrow purposes, such as when users are presented with very granular and specific purposes, overwhelm users by giving them too many options (see Figure 3 for an example). This is the case of the French DPA suggests several granular purposes in its guidelines [12]. Singh et al. have shown that users prefer having three purpose options consisting of *Required*, *Functional*, and *Advertising* cookies [60].

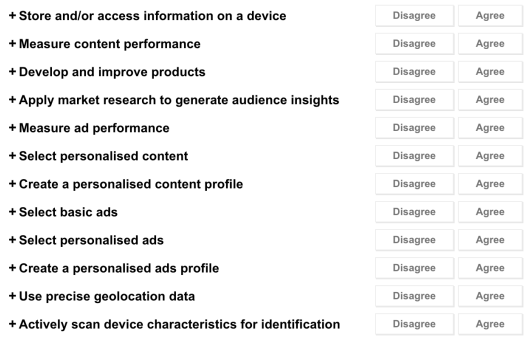

Figure 3: An example of a privacy notice with very specific purposes, taken from the wild.

Purpose names should be more specific about who they benefit. In the case of *Strictly Necessary / Essential / Required* purposes, some participants were skeptical about whether they were "necessary" for the user, or for the organization. As such, we recommend that organizations be clearer and more transparent about who these purposes benefit. For instance, this purpose could be renamed to *"Essential for basic website functions"*, or *"Required by the company"* to avoid potentially misleading users.

***Advertising* and *Personalized Advertising* purposes need to be better differentiated.** Almost all participants believed *Advertising* purposes, which are for non-targeted advertising, were the same as *Personalized Advertising* purposes. Therefore, we recommend that *Advertising* purposes clarify that this is for non-targeted advertising purposes to avoid confusing users. For example, the name could be changed to *"Non-Personalized Advertising"* to indicate the difference.

Purpose names should avoid conjunctions. Research in psychology has proven that when various items are grouped together, forming a conjunction, users will often remember the first and/or last items best, a phenomenon known as the *serial position effect* [52]. To mitigate the effects of this phenomenon, we recommend that data controllers only present one purpose to users at a time instead of grouping them together. As posited by Santos et al. [55], the use of bundling infringes upon the purpose specification principle.

5.1.1 Some purpose names were preferred over others. In our study, we found that participants often preferred one name over another in the case of purposes which were grouped together, as explained below. Therefore, we recommend that organizations consider using the preferred names instead of conjunctions for these purposes.

Statistics and Analytics. Participants preferred *Analytics* because it is more straightforward about what is happening to their data, whereas *Statistics* is more vague and sounds more technical. Participants felt saying "statistics" or "analytics" on its own without providing context about what organizations were collecting data for was misleading.

Performance and Functionality. Participants preferred *Performance* because it fit more with their perceptions of this purpose, believing it is meant to improve the site experience. *Functionality*, on the other hand, sounds like it is a purpose necessary for the website to function, which was deemed to be misleading.

Personalized Content The name was not intuitive for some participants who believed it meant the same as *Personalized Advertising* because "content" could also refer to advertising content. This misunderstanding could stem from the fact that most users are attuned to how organizations use data for delivering targeted ads [28, 40, 45]. Therefore, it is recommended that the name be modified to help users better differentiate it from *Personalized Advertising*, such as *Personalized [insert application, e.g., Video/Search] Recommendations*.

5.2 Towards User-Informed Purpose Descriptions

Descriptions are lacking crucial information users wish to know. Our findings suggest that purpose descriptions need to be more transparent about what organizations are doing with users' data. Participants indicated wanting to know more about i) how long their data will be retained for, ii) how to go about deleting their data and reversing their previous consent decisions, and iii) the sensitivity of the information organizations are collecting from users.

When descriptions mentioned they were personalizing ads and content or creating user profiles, many participants wanted more transparency from organizations about how their profiles were built and used. This is especially important as user data is often used to train AI models [2], and some organizations, such as Zoom, are updating their Terms of Service and not allowing users to refuse data collection for training AI models [29, 51].

Users want to know how their data is being used by organizations, and therefore more transparency about how personalization works needs to be conveyed to users. A UK Data Protection Authority (DPA) report found that when participants were informed about how the adtech system worked, participants were less likely to accept seeing online advertisements, revealing how being informed can change users' attitudes towards data sharing practices [63].

We denote that the law mandates information disclosures about the risks and consequences of processing purposes. Under the legal requirement of *informed* consent and the transparency principle, personal data processing must be handled in a transparent manner in relation to the user (Article 5(1)(a)), including obligations for websites to inform users about the types of data processes, data recipients (Article 14), the scope, consequences, [20] and *risks* in relation to the processing of personal data (Recital 39). Offering users legal information for consent to tracking empowers them, but also may induce negative impacts, as receiving such information may decrease one's perception of risk [61].

Our study upholds these legal requirements as we found that users also want to have such information. Both the law and user's intentions are aligned in the sense that both assume an informed and rational user that acts deliberately towards privacy-friendly options. In practice, users may not read such information and act against their own intentions, presenting a *privacy paradox*, where users' intentions do not match their behaviours [4].

Descriptions should avoid using loss aversion language. Participants preferred when purpose descriptions gave examples of services the purpose provides, but sometimes found it threatening when descriptions used loss aversion language by saying they would miss out on certain services by denying cookies. When information is framed negatively, it may put pressure on users by exploiting loss aversion [1] and nudge them towards consenting [55], especially when it is unclear which functionalities will be lost.

Negative framing may nudge users towards the website's wishes when the information about what is missing is omitted, hence violating a freely given consent requirement. Bongard et al. showed that users may develop incorrect mental models of the consequences of (not) consenting to data collection and processing [10]. We suggest that instead of using loss aversion language, descriptions could instead mention the services that are still provided if users reject their consent.

Participants want reassurance from organizations. Another commonly occurring theme within our findings are that participants want *reassurance* from organizations in their purpose descriptions. For example, participants said this can include providing them with more information to find out more about a technical aspect of data collection if they wanted (through a link or expandable definition), or more commonly, they wanted reassurance from organizations that their data would be kept private and secure, such as how it be anonymized, would not be used for profiling, and not shared with third parties, which is a finding in line with previous research [44].

However, we caution the use of reassurance in purpose descriptions; in some cases, reassurance is only a half-truth, such as in the case of *Marketing and Advertising* where one description said "We and third party companies / our partners use trackers for the purpose of measuring the audience of advertising on the site or application, without profiling you." In this case, it is true that the service provider (website) is not profiling users, but the third parties that data is being sent to are indeed profiling users, making these descriptions not fully accurate to users [5, 27]. Therefore, descriptions need to be careful in presenting accurate information when they wish to reassure users.

5.3 Improving Purposes: a Step Towards Improving the Consent Ecosystem

This study focused on users' perceptions of data collection purposes and descriptions that are commonly used in consent notices. We suggest that more focus be paid on combining our recommendations along with other broader consent ecosystem changes to further improve informed consent. Deceptive practices, such as only changing the purpose names and descriptions without changing the underlying consent ecosystem or data collection procedures, can undermine efforts towards truly informed consent. Ideally, changes to the online consent ecosystem would combine user-informed recommendations, along with appropriate technical and legal measures to prevent deceptive practices from taking advantage of users.

We foresee that informed purposes can be achieved through two methods: by implementing a consent nutrition label, and by adopting informed consent practices from pre-existing consent applications to further improve the *informed* aspects of informed consent.

5.3.1 Consent Nutrition Label. To account for the lack of user attention on security and privacy, researchers have tried implementing other ways of conveying this information [17, 37, 38]. Privacy

policies, which are often considered to be difficult and boring to read [58], have been improved through the use of privacy nutrition labels which use a tabular format to enhance user comprehension of an organization's privacy policy. They have been shown to increase information finding, and allow for users to make better comparisons between policies [37]. By leveraging better UI design, icons, and colours, users are more likely to pay attention to privacy policies, preferring them over traditional privacy policies, thus making privacy policies more accessible to users [37, 38].

A finding from our study was that, similar to privacy policies, few participants read consent notices, finding them to be boring, difficult to read, and annoying, confirming previous studies [43, 65]. Therefore, a possible use case for privacy nutrition labels is to apply them to the description of purposes within consent notices to make them more accessible and enjoyable to interact with. The suggestion of leveraging design elements in consent notices is aligned with the GDPR (Article 12(7)) which prescribes that information disclosed to data subjects may be provided in combination with standardised icons in order to give a meaningful overview of the intended processing in an easily visible, intelligible, and clearly legible manner.

5.3.2 Learning from Other Consent Applications. The current state of data collection purposes and their descriptions contribute to making online consent largely uninformed [24, 49, 53, 65]. We posit that much can be learned from other, better-established areas where consent and data collection information is conveyed, such as in human subject research and healthcare settings. In these fields, informed consent is conveyed to users in a variety of formats which might translate effectively to consent notices, as supported by Andreotta et al. [2].

Human subject research. In human subject research, there is often an ethics review board (ERB) that requires researchers specify the study's data handling procedures in the ethics form, in addition to a consent form that participants are required to read and sign. Data handling procedures are specific and detailed, and often listed in the consent form and/or orally conveyed to participants.

Our study found that purpose descriptions are lacking in crucial information that most participants wanted to know more about. As such, purpose descriptions could learn from the ethics review process for human subject research and include information that participants want to know, such as how long their data will be retained for, how users can request their data be deleted, and who will have access to user data.

Healthcare. In healthcare settings, patients are informed about the risks, benefits, and alternatives for a medical procedure [2, 54]. Similarly, in online consent, users should be made aware of the risks, benefits, and alternatives for consent to an organization's consent terms. Currently, users are often only presented with the benefits when they consent [45, 55] or the negatives of rejecting consent [6, 28, 48].

5.4 Limitations

We only conducted our interviews with fluent English speakers who were living in the UK and Ireland to ensure exposure to English-language privacy notices. Therefore, there may be linguistic nuances and other differences that might be present in privacy notices written in other languages. As we only focused on commonly-used data collection purposes used in the EU/UK, it is also possible that there are other purposes and language nuances presented in other jurisdictions we did not capture in this study. Hence, we do not generalize our findings to non-English languages or other data privacy laws due to these potential differences.

On a related note, privacy notices and the law are constantly changing [14]. As such, we may not be able to capture all the recent changes in the data collection purposes or descriptions we showed participants, nor all the purpose name and description variations used in privacy notices.

Our paper does not fully address the complex adtech and multi-actor nature underlying privacy notices. We conducted our study on users' perceptions of what is presented in consent notices, because this is what users are seeing first-hand. Consent notices are one of the only ways in which information about how user data may be processed is disclosed to users. Our findings contribute to the field of HCI, but a transdisciplinary solution that involves relevant stakeholders in the consent ecosystem is necessary to enact change.

Lastly, participants' perceptions of, and suggestions for improving purposes and their descriptions may not line up with how they would actually act in real life, a phenomenon called the *privacy paradox* [4]. This is why designing privacy notices void of deceptive designs and using user-friendly language is important to prevent users from choosing the easiest, deceptive options. We conducted a qualitative study to understand user perceptions of data collection purposes, laying the groundwork for future quantitative research.

5.5 Future Directions

Building upon our work, a controlled lab study looking into the efficacy of our proposed solutions, such as the consent nutrition label, applying consent mechanisms from other domains, and our suggestions for better purpose names and descriptions can yield insights into what the future of informed consent could look like. We do not measure how perceptions impact user behaviours, so a follow-up study investigating how perceptions may differ from behaviours could bring important insights. Additionally, since our study only looked at English privacy notices, future work should expand upon this by studying privacy notices in other languages to suggest improvements for non-English banners.

6 CONCLUSION

Through semi-structured interviews with 23 UK and Ireland-based participants, we studied user perceptions of data collection purposes, which form the basis of what users are consenting to share their data for. We investigated how six common purposes (*Strictly Necessary / Essential / Required, Statistics and Analytics, Performance and Functionality, Advertising, Personalized Advertising,* and *Personalized Content*) were perceived by users, and identified elements of an effective purpose name and description.

Our results suggest that most purpose descriptions were not informative enough, according to participants. Descriptions do not often tell users the specifics about an organization's data handling procedures, such as how long their data is retained, nor outline the data deletion process. Overall, participants wanted descriptions to provide more transparency and reassurance.

For purpose names, some names were preferred over others because they were perceived to be transparent or easy to understand for participants. It was common for participants to get some purposes confused with each other, or believe all purposes were covertly being used for advertising purposes. From our findings, we provide suggestions for how purpose names and descriptions can be improved, and envision a future of informed consent that is more user-centred.

acmart

REFERENCES

[1] Alessandro Acquisti, Manya Sleeper, Yang Wang, Shomir Wilson, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, and et al. 2017. Nudges for privacy and security: Understanding and assisting users' choices online. *Comput. Surveys* 50, 3 (Aug 2017), 1–41. https://doi.org/10.1145/3054926

[2] Adam J Andreotta, Nin Kirkham, and Marco Rizzi. 2022. AI, big data, and the future of consent. *AI & Society* 37, 4 (2022), 1715–1728.

[3] Manu Bansal. 2021. Flying Blind: How Bad Data Undermines Business. https://www.forbes.com/sites/forbestechcouncil/2021/10/14/flying-blind-how-bad-data-undermines-business/

[4] Susanne Barth and Menno DT De Jong. 2017. The privacy paradox: Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics* 34, 7 (2017), 1038–1058.

[5] Robert Bateman. 2023. Do I Need a Privacy Policy if I Don't Collect Any Data? Available at https://www.termsfeed.com/blog/privacy-policy-no-data-collected/.

[6] Benjamin Maximilian Berens, Heike Dietmann, Chiara Krisam, Oksana Kulyk, and Melanie Volkamer. 2022. Cookie disclaimers: Impact of design and users' attitude. In *Proceedings of the 17th International Conference on Availability, Reliability and Security*. 1–20.

[7] Asia J Biega. 2023. Data Repurposing through Compatibility: A Computational Perspective. *Journal of Institutional and Theoretical Economics* (2023).

[8] European Data Protection Board. 2007. Opinion 4/2007 on the concept of personal data (WP 136), adopted on 20.06.2007. https://ec.europa.eu/justice/article-29/documentation/opinionrecommendation/files/2007/wp136_en.pdf

[9] Dino Bollinger, Karel Kubicek, Carlos Cotrini, and David Basin. 2022. Automating Cookie Consent and GDPR Violation Detection. In *31st USENIX Security Symposium (USENIX Security 22)*. 2893–2910.

[10] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. 2021. "I am definitely manipulated, even when I am aware of it. It's ridiculous!" - Dark Patterns from the End-User Perspective. *Proceedings of ACM DIS Conference on Designing Interactive Systems* (2021). https://doi.org/10.1145/3461778.3462086

[11] Elijah Robert Bouma-Sims, Megan Li, Yanzi Lin, Adia Sakura-Lemessy, Alexandra Nisenoff, Ellie Young, Eleanor Birrell, Lorrie Faith Cranor, and Hana Habib. 2023. A US-UK Usability Evaluation of Consent Management Platform Cookie Consent Interface Design on Desktop and Mobile. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–36.

[12] Commission Nationale de L'informatique. 2020. Cookies et autres traceurs. Available at https://www.cnil.fr/fr/cookies-et-autres-traceurs/regles/cookies/lignes-directrices-modificatives-et-recommandation.

[13] Commission Nationale de L'informatique. 2022. Cookies: the CNIL fines Google a total of 150 million euros and Facebook 60 million euros for non-compliance with French legislation. https://www.cnil.fr/en/cookies-cnil-fines-google-total-150-million-euros-and-facebook-60-million-euros-non-compliance

[14] Martin Degeling, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz. 2019. We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. In *Proceedings 2019 Network and Distributed System Security Symposium*. Internet Society. https://doi.org/10.14722/ndss.2019.23378

[15] European Data Protection Board (EDPB). 2013. Opinion 03/2013 on purpose limitation (WP 203). Available at https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.

[16] European Data Protection Board (EDPB). 2013. Working Document 02/2013 providing guidance on obtaining consent for cookies, adopted on 2 October 2013. Available at https://www.pdpjournals.com/docs/88135.pdf.

[17] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the experts: What should be on an IoT privacy and security label?. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 447–464.

[18] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. 2019. Exploring how privacy and security factor into IoT device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[19] European Commission. 2018. 2018 Reform of EU data protection rules. Available at https://ec.europa.eu/commission/sites/beta-political/files/data-protection-factsheet-changes_en.pdf.

[20] European Commission. 2018. Guidelines on transparency under Regulation 2016/679, WP260 rev.01. Available at https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227.

[21] European Data Protection Board. 2023. *Guidelines 3/2022 on Deceptive Patterns in Social Media Interfaces: How to recognise and avoid them*. Technical Report Version 2.0. https://edpb.europa.eu/system/files/2023-02/edpb_03-2022_guidelines_on_deceptive_design_patterns_in_social_media_platform_interfaces_v2_en_0.pdf

[22] European Data Protection Board. 2023. Report of the work undertaken by the Cookie Banner Taskforce . Available at https://edpb.europa.eu/system/files/2023-01/edpb_20230118_report_cookie_banner_taskforce_en.pdf.

[23] European Parliament. 2002. Privacy and Electronic Communications Directive (ePrivacy Directive).

[24] Anne Josephine Flanagan, Jen King, and Sheila Warren. 2020. Redesigning data privacy: Reimagining notice & consent for human-technology interaction. Available at https://www3.weforum.org/docs/WEF_Redesigning_Data_Privacy_Report_2020.pdf.

[25] Imane Fouad, Cristiana Santos, Feras Al Kassar, Nataliia Bielova, and Stefano Calzavara. 2020. On Compliance of Cookie Purposes with the Purpose Specification Principle. In *International Workshop on Privacy Engineering (IWPE 2020)*. Genova, Italy, 1–8. https://hal.inria.fr/hal-02567022

[26] Barney G Glaser and Anselm L Strauss. 2017. *Discovery of grounded theory: Strategies for qualitative research*. Routledge.

[27] Google Developers. 2023. Third Parties. Available at https://web.dev/learn/privacy/third-parties/.

[28] Hana Habib, Megan Li, Ellie Young, and Lorrie Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–27.

[29] Smita Hashim. 2023. How Zoom's terms of service and practices apply to AI features. Available at https://blog.zoom.us/zooms-term-service-ai/.

[30] Monique Hennink and Bonnie N Kaiser. 2022. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social science & medicine* 292 (2022), 114523.

[31] Maximilian Hils, Daniel W Woods, and Rainer Böhme. 2020. Measuring the emergence of consent management on the web. In *Proceedings of the ACM Internet Measurement Conference*. 317–332.

[32] Kwesi Hughes-Lartey, Meng Li, Francis E Botchey, and Zhen Qin. 2021. Human factor, a critical weak point in the information security of an organization's Internet of things. *Heliyon* 7, 3 (2021), e06522.

[33] Information Commissioner's Office. 2019. Guidance on the use of cookies and similar technologies. Available at https://ico.org.uk/for-organisations/direct-marketing-and-privacy-and-electronic-communications/guide-to-pecr/guidance-on-the-use-of-cookies-and-similar-technologies/.

[34] Information Commissioner's Office. 2023. Cookies and similar technologies. Available at https://ico.org.uk/for-organisations/direct-marketing-and-privacy-and-electronic-communications/guide-to-pecr/cookies-and-similar-technologies/.

[35] Information Commissioner's Office. 2023. Overview – Data Protection and the EU. Available at https://ico.org.uk/for-organisations/data-protection-and-the-eu/overview-data-protection-and-the-eu/.

[36] Italian Data Protection Authority. 2021. Guidelines on Cookies and Other Tracking Tools. Available at https://www.garanteprivacy.it/documents/10160/0/GUIDELINES+ON+COOKIES+AND+OTHER+TRACKING+TOOLS+-+Executive+Summary.pdf/2278a9d4-a0e1-1578-1a09-8291106f4591?version=3.0.

[37] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. 2009. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*. 1–12.

[38] Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. 2010. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1573–1582.

[39] Stefan Korff and Rainer Böhme. 2014. Too much choice: End-user privacy decisions in the context of choice proliferation. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. 69–87.

[40] Anastasia Kozyreva, Philipp Lorenz-Spreen, Ralph Hertwig, Stephan Lewandowsky, and Stefan M Herzog. 2021. Public attitudes towards algorithmic personalization and use of personal data online: Evidence from Germany, Great Britain, and the United States. *Humanities and Social Sciences Communications* 8, 1 (2021), 1–11.

[41] Chiara Krisam, Heike Dietmann, Melanie Volkamer, and Oksana Kulyk. 2021. Dark patterns in the wild: Review of cookie disclaimer designs on top 500 German websites. In *Proceedings of the 2021 European Symposium on Usable Security*. 1–8.

[42] Oksana Kulyk, Nina Gerber, Annika Hilt, and Melanie Volkamer. 2020. Has the GDPR hype affected users' reaction to cookie disclaimers? *Journal of Cybersecurity* 6, 1 (2020), tyaa022.

[43] Oksana Kulyk, Annika Hilt, Nina Gerber, and Melanie Volkamer. 2018. "This website uses cookies": Users' perceptions and reactions to the cookie disclaimer. In *European Workshop on Usable Security (EuroUSEC)*, Vol. 4.
[44] Oksana Kulyk, Karen Renaud, and Stefan Costica. 2023. People want reassurance when making privacy-related decisions—Not technicalities. *Journal of Systems and Software* 200 (2023), 111620.
[45] Lin Kyi, Sushil Ammanaghatta Shivakumar, Cristiana Teixeira Santos, Franziska Roesner, Frederike Zufall, and Asia J Biega. 2023. Investigating deceptive design in GDPR's legitimate interest. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
[46] Latvian Data Protection Authority (DVI). 2022. Guidelines for the use of cookies on the website. Available at https://www.dvi.gov.lv/lv/jaunums/par-preventivas-parbaudes-attieciba-uz-sikdatnu-izmantosanas-atbilstibu-lielako-latvijas-e-komersantu-timekla-vietnes-rezultatiem.
[47] Latvian DPA (DVI). 2022. Guidelines for the use of cookies on the website. Available at https://media.licdn.com/dms/document/C4D1FAQGpDp65vdTl_g/feedshare-document-pdf-analyzed/0/1649147267223?e=1678924800&v=beta&t=7MXnitOQ9-bf1hXT092TRmS9aQBfLMjkGe2J0ABR7gk.
[48] Eryn Ma and Eleanor Birrell. 2022. Prospective consent: The effect of framing on cookie consent decisions. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–6.
[49] Dominique Machuletz and Rainer Böhme. 2020. Multiple purposes, multiple problems: A user study of consent dialogs after GDPR. *Proceedings on Privacy Enhancing Technologies (PoPETS)* (2020).
[50] Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica* 22, 3 (2012), 276–282.
[51] Matt Milano. 2023. Zoom Updates Terms to Use Customer Data for AI Training With No Opt-Out. Available at https://www.webpronews.com/zoom-updates-terms-to-use-customer-data-for-ai-training-with-no-opt-out/.
[52] Bennet B Murdock Jr. 1962. The serial position effect of free recall. *Journal of Experimental Psychology* 64, 5 (1962), 482.
[53] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[54] Parth Shah, Imani Thornton, Danielle Turrin, and John E Hipskind. 2022. Informed Consent. Available at https://www.ncbi.nlm.nih.gov/books/NBK430827/.
[55] Cristiana Santos, Arianna Rossi, Lorena Sanchez Chamorro, Kerstin Bongard-Blanchy, and Ruba Abu-Salma. 2021. Cookie banners, what's the purpose? Analyzing cookie banner text through a legal lens. In *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society*. 187–194.
[56] Martina Angela Sasse, Sacha Brostoff, and Dirk Weirich. 2001. Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. *BT Technology Journal* 19, 3 (2001), 122–131.
[57] Benjamin Saunders, Julius Sim, Tom Kingstone, Shula Baker, Jackie Waterfield, Bernadette Bartlam, Heather Burroughs, and Clare Jinks. 2018. Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & quantity* 52 (2018), 1893–1907.
[58] Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. 2015. A design space for effective privacy notices. In *Eleventh Symposium on Usable Privacy and Security (SOUPS 2015)*. 1–17.
[59] Fatemeh Shirazi and Melanie Volkamer. 2014. What deters Jane from preventing identification and tracking on the web?. In *Proceedings of the 13th Workshop on Privacy in the Electronic Society* (Scottsdale, Arizona, USA) *(WPES '14)*. Association for Computing Machinery, New York, NY, USA, 107–116. https://doi.org/10.1145/2665943.2665963
[60] Ashutosh Kumar Singh, Nisarg Upadhyaya, Arka Seth, Xuehui Hu, Nishanth Sastry, and Mainack Mondal. 2022. What cookie consent notices do users prefer: A study in the wild. In *Proceedings of the 2022 European Symposium on Usable Security*. 28–39.
[61] Joanna Strycharz, Edith Smit, Natali Helberger, and Guda van Noort. 2021. No to cookies: Empowering impact of technical and legal knowledge on rejecting tracking cookies. *Computers in Human Behavior* 120 (2021), 106750. https://doi.org/10.1016/j.chb.2021.106750
[62] Jon Swain. 2018. A hybrid approach to thematic analysis in qualitative research: Using a practical example. *Sage Research Methods* (2018).
[63] UK Information Commissioner's Office. 2019. Adtech: Market Research Report. Available at https://ico.org.uk/media/about-the-ico/documents/2614568/ico-ofcom-adtech-research-20190320.pdf.
[64] UK Legislation. 2018. Data Protection Act 2018. https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted
[65] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un)informed consent: Studying GDPR consent notices in the field. In *Proceedings of the 2019 ACM SIGSAC conference on computer and communications security*. 973–990.
[66] Michael Veale, Midas Nouwens, and Cristiana Santos. 2022. Impossible asks: can the transparency and consent framework ever authorise real-time bidding after the Belgian DPA decision? *Michael Veale, Midas Nouwens and Cristiana Teixeiras Santos, Impossible asks: Can the Transparency and Consent Framework Ever Authorise Real-Time Bidding After the Belgian DPA Decision* (2022), 12–22.