Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence

Midas Nouwens[1,2] {midasnouwens}
Ilaria Liccardi[2] {ilaria}
Michael Veale[3] {m.veale}
David Karger[2] {karger}
Lalana Kagal[2] {lkagal}

[1]{}@cavi.au.dk
Digital Design & Information Studies
Aarhus University, DK

[2]{}csail.mit.edu
MIT CSAIL
Cambridge, MA, USA

[3]{}ucl.ac.uk
Faculty of Laws
UCL, UK

ABSTRACT

New consent management platforms (CMPs) have been introduced to the web to conform with the EU's General Data Protection Regulation, particularly its requirements for consent when companies collect and process users' personal data. This work analyses how the most prevalent CMP designs affect people's consent choices. We scraped the designs of the five most popular CMPs on the top 10,000 websites in the UK (n=680). We found that dark patterns and implied consent are ubiquitous; only 11.8% meet our minimal requirements based on European law. Second, we conducted a field experiment with 40 participants to investigate how the eight most common designs affect consent choices. We found that notification style (banner or barrier) has no effect; removing the opt-out button from the first page increases consent by 22–23 percentage points; and providing more granular controls on the first page decreases consent by 8–20 percentage points. This study provides an empirical basis for the necessary regulatory action to enforce the GDPR, in particular the possibility of focusing on the centralised, third-party CMP services as an effective way to increase compliance.

Author Keywords

Notice and Consent; Dark patterns; Consent Management Platforms; GDPR; Web scraper; Controlled experiment

CCS Concepts

•**Information systems** → **Online advertising;** •**Security and privacy** → **Usability in security and privacy;** •**Social and professional topics** → **Privacy policies;** •**Applied computing** → **Law;**

INTRODUCTION

The predominant method of giving people some semblance of control over their privacy while browsing the web is 'notice and choice' or 'notice and consent' [20] . These mechanisms involve showing an individual an informational statement and, depending on their (in)action, acquiring or assuming their agreement to collecting, storing, and processing their data. To many, this practice has become informally known as 'cookie banners'.

What counts as sufficient notice, and what counts as legally-acceptable consent, significantly differs depending on the geographical and regulatory scope that an actor falls in. The application in Europe of the General Data Protection Regulation (GDPR) [26] from May 2018, together with recent regulatory guidance from data protection authorities (DPAs) and jurisprudence from the Court of Justice of the European Union (CJEU), has highlighted the illegality of the way 'notice and consent' has hitherto functioned in the EU. These regulatory changes have both clarified the concept of consent in European law, as well as brought more significant (and extraterritorial) consequences for flaunting these rules. EU law in particular focuses on the *quality* of the consent required, and its freely-given, optional nature.

Consent management platforms (CMPs) have gained traction on the Web to help website owners outsource regulatory compliance. These (often third-party) code libraries purport to help websites establish a lawful basis to both read and write information to users' browsers and to process these individuals' personal data, often for the purposes of tracking and complex advertising transactions, such as 'real-time bidding' [31].

This intertwining of interface designs and data protection and privacy law raises significant questions. This paper deals with two of them:

1. What is the current state of interface design of CMPs in the EU, and how prevalent are non-compliant design elements?
2. How do interface designs affect consent actions of users and, by extension, how 'freely given' that consent is?

To answer the first question, we surveyed the designs of the 5 most commonly used third-party CMPs by scraping their varied implementations on the top 10,000 most popular websites in the United Kingdom (UK) (n=680); and evaluated them against European law and regulatory guidance. To answer the second question, we built a browser plugin that injects consent notices into webpages and ran a controlled experiment (n=40) with eight different interfaces to see how they affect participants' consent responses.

CONSENT AND WEB TECHNOLOGIES UNDER EU LAW

EU law considers users' devices and information within them part of their private sphere. Relevant protection is extended to all EU residents and to all individuals around the world being delivered

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376321

online services from the Union. In light of a growing trend in the early 2000s of rightsholders sneaking piracy-spotting rootkits onto users' devices [35], the ePrivacy Directive [25] was amended to require that storing or accessing information on a user's device not 'strictly necessary' for providing an explicitly requested service requires both clear and comprehensive information and opt-in consent [35]. This also applies to cookies, HTML web storage, and fingerprinting in browsers providing non-essential features such as tracking. Such consent is however, *not* required for essential functions such as remembering login status, a shopping cart, or cookies for data security required by law [14, 30].

The ePrivacy Directive is connected to definitions in European data protection law, so when the GDPR [26] repealed and replaced the Data Protection Directive 1995 [24] in 2018, these practices became subject to new, heightened standards concerning the quality of consent. The GDPR defines consent as "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her" [26, art 4(11)]. Several aspects of this legal regime with design implications are important to highlight here, which are drawn from the legal texts, regulators' guidance, and court cases or opinions.

Freely given and unambiguous consent

Regulators and the Court of Justice of the EU have both emphasised that for consent to be freely given and informed, it must be a separate action from the activity the user is pursuing [3, 6, 18]. So-called 'implicit' or 'opt-out' consent — continuing to use a website without active objection to a notice — is not a clear positive action and as such will not establish a valid legal basis to lay cookies or process data on the basis of consent [6, 14, 30].

As a consequence of the importance of the freely given nature of consent, design matters for legal compliance. Pre-ticked boxes, which require a positive action to opt-out from, are explicitly singled out in the GDPR as an invalid form of consent [26, recital 32]. The Court of Justice has recently ruled that they were also not a valid form of consent under the previous law, operational since the mid-90s [18]. The UK's Information Commissioner's Office further states that "[a] consent mechanism that emphasises 'agree' or 'allow' over 'reject' or 'block' represents a non-compliant approach, as the online service is influencing users towards the 'accept' option." Similarly, cookie boxes without a 'reject' option, or where it is located in a 'more information' section or on a third party webpage, are also non-compliant [30]. One of the CJEU's Advocates General (official impartial advisors to the Court on cases raising new points of law) has emphasised the need that both actions, "optically in particular, be presented on an equal footing" [3, para 66].

Moreover, it must be "as easy to withdraw as to give consent" [26, art 7(3)]. This means if consent was gathered through "only one mouse-click, swipe of keystroke", withdrawal must be "equally as easy" and "without detriment" or "lowering service levels" [6].

An issue of continued contention is the validity of so-called 'cookie walls', whereby consent is a prerequisite to accessing a website. While several regulators appear minded to suggest in many or all cases this practice is illegal [7, 14, 23, 30], the issue remains unclear [56] and the final conclusion will regardless be subject to the "glacial flow" of the draft ePrivacy Regulation through the EU's legislative process [51].

Specific and informed consent

An important aspect of data protection is *purpose limitation*, meaning users must consent in relation to a particular and specific purpose for processing data. They cannot provide *carte blanche* for a data controller to do whatever they like.[1] These purposes cannot be inextricably 'bundled', so an 'accept all' button is only compliant if it is additional to the possibility of specifically consenting to each purpose [14].

Furthermore, consent is invalid unless all organisations processing this data are specifically named [14, 31]. Simply linking to an external list of potential vendors, which may not represent the code being run on the linking webpage, is "insufficient to provide for free and informed consent" [31]. Consent should be able to be rejected at the same level as the 'accept' button, so having to navigate further to third party websites to reject tracking is non-compliant [30]. Information required to be provided to data subjects includes certain GDPR–mandated information (including controller contact, processing purposes, legal basis, recipients and sources of data, international transfers, storage period, data rights and rights of complaint, and meaningful information about the logic of significant automated decision-making) [26, arts 13–14], as well as the duration of cookies [3, 30].

Efficient and timely data protection

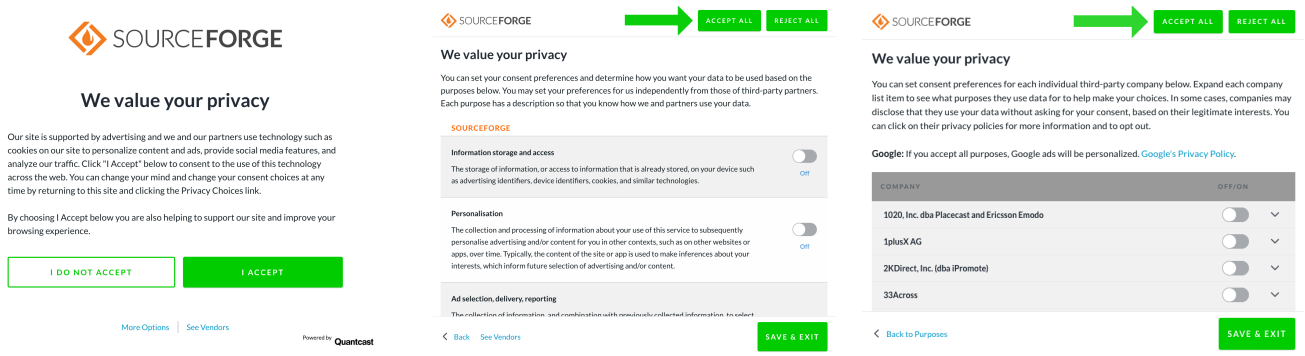
Individuals have the right to 'efficient and timely' protection of their data rights, meaning where consent is required, it is required prior to data processing, not subsequently [6, 17]. Cookies must not be set before the user has expressed their affirmative consent. Furthermore, fresh consent is required when new, non-essential cookies are being set by a new third party [6, 30]. The burden is on the data controller to be able to demonstrate that they adhere to data protection law and principles, including that they have valid consent for each individual [26, art 5(2)].

RELATED WORK

Notice & Consent

The predominant model for communicating information privacy protections to end-users has been notice(/awareness) and consent(/choice). The interface designs of this model have mostly been privacy policies and opt-in/out interfaces [20], which legally can be seen as "pre-formulated declarations of consent", or "click-wrap" contracts [13]. The usability challenges of these interfaces have seen considerable work across disciplines, largely divisible in studies that establish the shortcomings of interface designs, and studies proposing alternative technologies. Privacy policy notices are notorious for taking a disproportionate amount of time to go through and require reading comprehension abilities at university level [32]. Privacy policies are rarely read by users [42, 43, 55] prior to using or visiting a site/service. Users have been shown to (almost automatically) consent without viewing them [2, 5, 8, 41, 42] since they stand in the way of the users' primary goal: accessing the service [2, 5]. This behaviour has been attributed to

[1] As an unalienable fundamental right, it is impossible for an EU resident to 'sign away' their right to effective data protection.



(a) First page **(b) Categories and purposes** **(c) Vendors/third-parties**

Figure 1. The three components of the QuantCast CMP on https://sourceforge.net as of September 2019.

the users' difficulty understanding how to make meaningful decisions about their privacy preferences; but even in situations where they are made aware of the implications of their decision, they prefer short-term benefits over long-term privacy [2]. Because of this, control mechanisms of these notices are considered illusory in practice [12] — sometimes having devolved into merely an informational statement rather than an interactive control panel.

The perceived ineffectiveness of this approach has given rise to a number of design alternatives (for an overview of the entire design space, see [46]). Gage Kelley et al. proposed standardised "nutrition label" notices with icons representing the type of data collected and how it is used, and showed how it helped users find information more quickly and accurately [34]. Reeder et al. developed an interactive matrix visualisation called Expandable Grid which shows a colour-coded overview of a policy that can be expanded for more detail [44]. The Platform for Privacy Preferences (P3P) was an involved attempt to help automate some of this process by building a machine-readable language for expressing website privacy policies which could then interface with user agents, such as the browser or other privacy applications [19]. While it was implemented by Microsoft for Internet Explorer and Edge, P3P never achieved widespread adoption, partly because its comprehensiveness was seen as too complex for regular website owners to apply but also because there was no regulatory or political impetus to force browser vendors to use it.

The majority of studies around notice and consent have focused on how well the interface design helps users make informed decisions. This paper focuses more on the legal *quality* of the consent that is collected.

Dark patterns

Interface designs that try to guide end-users into desired behaviour through malicious interaction flows are referred to as "dark patterns" [29]. As a phenomenon they are part of the larger research agenda around persuasive design [27] and nudging [1, 50]. The practice of dark patterns for privacy notices — while only sometimes discussed under this moniker in HCI and privacy literature [9, 15, 29, 38] — is extensively reported on by consumer protection organisations [28], white papers [49], and popular press [47] (for an excellent overview, see the Norwegian Forbrukerrådet document "Deceived by Design" [28]). Its infamy has led the European Union and data protection officers to specifically highlight certain common dark patterns as non-compliant examples of the GDPR in its advisory documents such as privacy intrusive default settings, hiding away privacy-friendly choices and requiring more effort from the user to select it, illusory or take-it-or-leave-it choices, etc. Senators from the United States have recently introduced a draft bill specifically aimed at outlawing such practices, stating that it should be prohibited for any large online operator to "design, modify, or manipulate a user interface with the purpose or substantial effect of obscuring, subverting, or impairing user autonomy, decision-making, or choice to obtain consent or user data" [21].

Since the submission of this paper, two studies have been released that look specifically at the consent management platforms that have appeared in response to the GDPR.

Utz et al [53] analysed a random sample of 1,000 CMPs and manually categorised them along different design dimensions (e.g., positioning, size, consent options). They found (among other things) that a minimum of 57.4% used dark patterns to nudge users to select privacy-unfriendly options, and that 95.8% provide either no consent choice or confirmation only. They conducted a follow-up experiment to test the effects of the CMP position, the granularity and nudging of choices, and the technicality of the language and presence of a privacy policy link. They demonstrate that positioning the CMP in the lower (left) part of the screen increases interaction rates; users are more likely to accept tracking given a binary choice than when given more granular options; acceptance rate increased from a mere 0.16% to 83.55% when options were preselected; and technical language and privacy policies have a minor effect on consent choice.

The work by Matte, Bielova, and Santos [39] investigates the actual consent signal sent from the CMP to the respective data processors. They detect that 12.3% of 1,426 sites send a consent signal before the user makes a choice. Semi-automatically reviewing 560 sites reveals that 54% of them contain at least one violation regarding the way consent is determined, asked, or complied with.

Empirical Studies of EU Privacy Regulation

Various studies have tried to chart the impact of European privacy regulation on the collection and processing of personal data on the web, both within its territorial scope and globally. A longitudinal 4-year study of the impact of the revised ePrivacy directive on

cookie placement shows that 1) 49% of websites placed cookies before receiving consent; 2) 28% of websites did not provide any consent mechanism; and 3) the percentage of websites violating the directive stayed constant over the course of 4 years, indicating the policy to be ineffective [52].

With respect to the GDPR, both industry and academia have been monitoring its effects since being introduced in May 2018. Degeling et al. [22] monitored the prevalence of privacy policies on websites before and after the introduction of the regulation, showing that in some EU member states the number of policies increased by 15.7% (to a total of 84.5%), while 72.6% of sites updated documents they already had. They estimate that a total of 62.1% of websites in Europe display a consent notice, an increase of 16% since shortly before the regulation became enforceable. Adzerk, an ad tech company, places this percentage considerably lower, at a mere 20.4% [4], although their methodology is more restrictive than Degeling et al.'s. Interestingly, QuantCast, one of the largest CMP providers, also brought out a report stating that over 90% of users (n=1bn) have consented to data processing [40].

Sanchez-Rola et al. [45] performed an evaluation of the tracking undertaken by 2,000 high-traffic websites and evaluated how information notices and actual tracking behaviour changed. They found that the GDPR affected EU and US sites in the same way, that consent management platforms reduced the amount of tracking, but that personal data collection is still ubiquitous: 90% still made use of cookies that were able to identify individual users. Sørensen and Sokol [48] present a more nuanced picture of the shifts in third-party tracker presence and behaviour, showing a decrease mostly present in private websites, whereas websites hosted by public institutions mostly stayed the same between February and September 2018. Along the same lines, there exists a difference between EU and non-EU private-sector sites, but little difference in public sites. Depending on the purpose category the tracker falls into, further distinctions can be made. The largest shift was visible in data collection for advertising, and the least in those used for cybersecurity. Overall, only 151 third-party trackers are used by 1% or more of the websites, while the remaining long-tail of 968 have a share of less than one percent.

STUDY 1: SCRAPING CMP INTERFACE DESIGNS

Little is known about many aspects of consent management platforms on the Web, particularly around the consent modalities, quality of this consent and related practices found in the field in the European Union. The major five CMP vendors offer a wide range of customisation options for their clients, and so from an identification of the CMP vendor it does not follow that many assumptions can be made about the interface design. To understand the status quo of consent management plaform interface design after the GDPR, we developed a web scraper to collect information about the five most commonly used third-party CMPs in the top 10,000 most-visited websites in the United Kingdom.

While their sophistication varies, surveyed CMPs all share similarities in back-end function. When a user accesses a site, the CMP detects their IP address and checks their cookies or local storage for any previously set consent preferences, and retrieves this data. If this fails, or if the CMP decides their preferences have expired, the user is shown a consent notice, and their response is recorded. This consent status is then passed on to any integrated tag firing rules, ad servers, and real-time bidding platforms the website has employed.

Visually, the CMP interfaces generally consist of three parts: 1) a first page describing the general purpose of the consent pop-up, with bulk consent options ('accept all' and, for some, 'reject all') (Fig. 1a); 2) a second page with a more detailed description of the different data processing categories or purposes (e.g. personalisation, marketing), the ability to toggle them individually or collectively, and a button to submit the current consent state (Fig. 1b); and, 3) a third page with a breakdown of all the vendors for whom the data is collected or with which it is shared, again with the ability to toggle individually or collectively, and a button to save these settings (Fig. 1c). Not all deployed CMPs have all parts of these interfaces enabled.

Method

We built a Web scraper to collect data about the CMP's visual elements, interaction design, and text content (e.g. names of data processing categories or vendors). The scraper utilised the Python library *Scrapy*[2] and JavaScript rendering service *Splash*[3]. The variables the scraper collected included the CMP vendor, the notification style (banner, barrier, other), the type of consent (explicit or implicit) and specific user actions counted as consent (consent/visit/navigation/reloading/scrolling/closing the pop-up/clicking the page); the existence of accept and reject all buttons and the minimum number of clicks to make them available; for both vendors and categories/purposes, the existence of lists of these, their extent and descriptions, whether or which are enabled for user control, and their default state(s).

We ran the scraper from a Danish IP address[4] over 3 days in September 2019 over the top 10,000 UK sites according to web-traffic service *Alexa*. We throttled our scraper to two concurrent URL requests and no concurrent requests per domain, with a delay of 2 seconds. We cycled through three different user agents copied from our browsers to make sure the websites treated us as normal visitors, rather than an automated crawler. The CMPs the scraper was designed for are third-party services as identified by *Adzerk* in August,[5] which together account for ~58% of the market share: QuantCast, OneTrust, TrustArc, Cookiebot, and Crownpeak. We targeted UK sites, rather than sites across all EU countries, because the Adzerk report gives us information about the total population of CMPs in the UK market. This allowed us to check that our scraper's sample was representative both in number of CMPs identified and the overall distribution of the five most popular ones.

To determine the presence of a particular CMP, the scraper looked for an identifying HTML element within 5–15 seconds of arriving on the site (depending on the particular CMP and how it injects the pop-up). Data to construct the variables were extracted by querying for elements and attributes, traversing the DOM if no unique indentifiers existed, or accessing globally scoped objects. This data was pushed to a *MongoDB* database. Before deployment, the data returned by the scraper was manually

[2] https://github.com/scrapy/scrapy

[3] https://github.com/scrapy-plugins/scrapy-splash

[4] Relevant legislation is harmonised across the EU and so a Danish IP and UK IP are the same jurisdiction for our purposes.

[5] A company that does server-side ad serving and writes reports about the state of the industry: www.adzerk.com

CMP	Sites	Median vendors (low./upp. quartiles)	Explicit/implicit consent	Banner/barrier	Preticked options	Minimum compliance
Cookiebot	12.5% (85)	104 (61, 232)	45/40	78/7	64 (75.3%)	2 (5.6%)
Crownpeak	12.2% (83)	38.5 (18.8, 132.3)	46/37	52/31	67 (80.7%)	0 (0%)
OneTrust	24.3% (165)	58 (26.5, 104.5)	47/118	158/7	108 (65.4%)	3 (1.8%)
QuantCast	41% (279)	542 (542, 542)	279/0	132/147	90 (32.3%)	73 (26.2%)
TrustArc	10% (68)	87 (38, 152)	42/26	26/42	53 (77.9%)	2 (2.9%)
all	**680**	**315 (58, 542)**	**459/221**	**446/234**	**382 (56.2%)**	**80 (11.8%)**

Table 1. Key statistics on scraped CMPs.

validated with 40 randomly selected sites from the list of 10,000 for each of the five CMPs. The scraper code and dataset will be available as supplementary material alongside the paper.

Understanding compliance

Based on the above section on EU law, we consider three core, measurable conditions that providers will have to meet to be considered legally compliant for the purpose of this study. This serves as a minimum hurdle: meeting these conditions alone will not guarantee compliance with the law, as there are a multitude of aspects and provisions, many of which can only be appropriately assessed qualitatively. However, these are conditions that are testable with the variables from our scraper, and therefore provide a window on the *maximum* level of compliance in the industry today. These conditions are:

Consent must be explicit This condition is true if consent is a clear, positive, affirmative act, such as clicking a button, rather than e.g. continuing to navigate a website.

Accepting all is as easy as rejecting all Consent must be as easy to give as to withdraw/refuse. This condition is met if accepting all takes the same number of clicks as rejecting all, and automatically not met in the case where consent requires no clicks (i.e. Condition 1 is violated)

No pre-ticked boxes Consent to any vendor or purpose must be through affirmative acts at all granularity. If no non-necessary purposes or vendors are automatically on, this condition is met.

Factors which could contribute to non-compliance which we did *not* examine include qualitatively considering the information provided (e.g. specificity of purposes, contact details of vendors, provision of the duration of cookies), nor certain visual features such as colour or size or prominence of buttons beyond clicks.

Results

680 (6.8%) of the top 10,000 UK websites contained a CMP which could be successfully scraped by our tool. According to a survey of the top 10K UK websites in August 2019 [4], only 20.35% of the top 10K UK websites are reported to use a CMP (from any vendor). 1191 of those (i.e., 58.52%) use the top 5 CMPs, which means the 680 instances our scraper captured represents 57.09% of the total population[6].

[6]It should be noted that Adzerk's methodology counts CMPs by URL endpoints of the Javascript files and we found during development that websites frequently include inactive CMPs' .js files. This means that Adzerk's statistics are likely inflated with double-counting, and that our survey is consequently more representative than the 57.09% would indicate.

We found that implicit consent is common among these sites (32.5%). An array of actions that websites count as consent (but which EU law does not) was extracted from their code, such as just visiting the site (16.8%), navigating within the site (6.2%), revisiting/refreshing the page (7.6%), scrolling or clicking on the page (5.3%) or closing the pop-up or banner (1.6%). 9% of sites accepted more than one form of implicit consent. With only a handful of idiosyncratic exemptions all implied consent was found in the use of 'banner' rather than 'barriers' (a barrier style is in Fig. 1). Within those CMPs exhibiting explicit consent, there was a roughly even split between the use of barriers and banners (50.3%/49.7%). Popular CMP implementation wizards still allow their clients to choose implied consent, even when they have already indicated the CMP should check whether the visitor's IP is within the geographical scope of the EU, which should be mutually exclusive. This raises significant questions over adherence with the concept of *data protection by design* in the GDPR.

The vast majority of CMPs make rejecting all tracking substantially more difficult than accepting it. 50.1% of sites did not have a 'reject all' button. Only 12.6% of sites had a 'reject all' button accessible with the same or fewer number of clicks as an 'accept all' button. In practice, this means both were accessible on the first page — an 'accept all' button was never buried in a second layer. 74.3% of reject all buttons were one layer deep, requiring two clicks to press; 0.9% of them were two layers away, requiring at minimum three.

Furthermore, when users went to amend specific consent settings rather than accept everything, they are often faced with pre-ticked boxes of the type specifically forbidden by the GDPR [26, recital 32]. 56.2% of sites pre-ticked optional vendors or purposes/categories, with 54.1% of sites pre-ticking optional purposes, 32.3% pre-ticking optional categories, and 30.3% pre-ticking both. Our scraper was detecting visual status rather than functional status—we do not know the impact on toggling on or off vendors or categories beyond what the CMP tells the user is happening (Matte et al.'s [39] findings indicate 7.7% of CMPs ignore the consent signal submitted by the user).

Sites relied on a large number of third party trackers, which would take a prohibitively long time for users to inform themselves about clearly. Out of the 85.4% of sites that did list vendors (e.g. third party trackers) within the CMP, there was a median number of 315 vendors (low. quartile 58, upp. quartile 542). Different CMP vendors have different average numbers of vendors, with the highest being QuantCast at 542 (see Table 1). 75% of sites had over 58 vendors. 76.47% of sites provide some descriptions

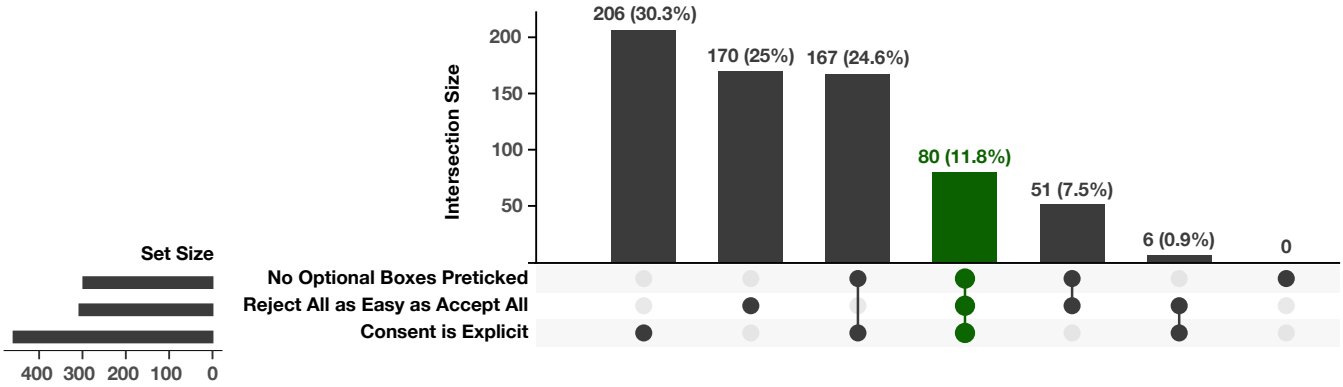


Figure 2. UpSet diagram [16, 36] of sites by adherence to three core conditions of EU law. Sites meeting all three in green.

of their vendors. The mean total length of these descriptions per site is 7985 words: roughly 31.9 minutes of reading for the average 250 words-per-minute reader, not counting interaction time to e.g. unfold collapsed boxes or navigating to and reading specific privacy policies of a vendor.

As discussed, we consider that a site is minimally compliant if it has no optional boxes pre-ticked, if rejection is as easy as acceptance, and if consent is explicit. Only 11.8% of sites met these basic requirements. The interaction between the requirements is shown in Figure 2. This varied significantly by CMP vendor — as shown in Table 1, only Quantcast has a non-negligible number of CMPs that we consider minimally compliant (26.2%), with Crownpeak having zero (that we found). This can largely be explained by the non-existence of implicit consent in QuantCast CMPs and their lower levels of pre-ticked boxes.

Interim Discussion

Given that all vendors (with the exception of Crownpeak) have examples in the wild of minimally compliant CMPs, it is unclear whether non-compliance is a practical result of sites configuring it in a non-compliant manner, being encouraged to do so by the CMP vendors or, in some cases, running older CMPs without updating them in light of the more publicised nature of the law.[7] Whatever the practical reasons, 11.8% is an extraordinarily low number for seemingly market-leading CMP vendors, and suggests an urgent role for data protection authorities to take action to ensure only correct configurations are permitted.

The dataset in this study will be available to other researchers, and we welcome further research into, for example, the scraped text content of the CMPs, as the 11.8% in this study is a maximum value that is likely to only decrease on consideration of further aspects of the law which are harder to assess in a formulaic manner.

Limitations

Although we manually validated the scraper, we cannot guarantee that there are no false negatives or false positives in our dataset. Because these CMPs are dynamically rendered via JavaScript, determining whether the state of the DOM scraped is the final one is tricky (further complicated by the fact that Scrapy's engine runs on ECMAScript 2015 making tools to deal with asynchronous execution, such as *async/await*, unavailable). We hardcoded a waiting time of 5-15 seconds between loading the site and scraping the content which should be more than sufficient, but there might be exceptions. The CMP might be customised either by the company or the website owner, thwarting the automated way we identify the presence of elements. Legacy implementations, either from various iterations over the years or because the company has been sold multiple times, also introduced branches in the CMP code we might have missed. While we did our best to identify and work around elements of the CMPs designed to obfuscate their function and prevent automation, deliberate changes to data retrieval are often used to foil research for those studying APIs [10, 11], and such practices seem likely in this domain also to protect against potential automated regulatory scrutiny.

[7]Note that the recent judgement from the European Court of Justice clarified that these requirements have been part of EU law since 2012, rather than just since the GDPR [18]

STUDY 2: EFFECTS OF DESIGNS ON ANSWERS

The goal of the second study was to establish if, and to what extent, certain CMP designs affect the consent answer given by users. We were interested in non-compliant designs that are very prevalent, or designs that are not yet described as non-compliant by the applicable regulation. We conducted two field experiments to establish the effects on user behaviour and consent rate of 1) barrier and banner notifications; 2) equal and unequal prominence of accept all and reject all options on the first page; and 3) the level granularity of consent options on the first page (bulk, purposes, vendors).

Method

Design

The study consisted of two counter-balanced experiments, evaluating a total of 8 different interfaces (see Figure 3).

Experiment 1 used a [2x2] latin square, within-subjects, repeated measures design. The independent variables were the NOTIFICATION STYLE (*Barrier; Banner*) and BULK CONSENT BUTTONS (*Accept all+Reject all; Accept all*). The primary dependent variable was the CONSENT ANSWER (*Accept all; Reject all; Submit default; Submit personalised*).

Experiment 2 used a [1x4] latin square, within-subjects, repeated measures design. The independent variable was the CONSENT GRANULARITY (*Bulk; Bulk+Purposes; Bulk+Vendors;*

Bulk+Purposes+Vendors) on the first page of the notification. The primary dependent variable was the CONSENT ANSWER (*Accept all; Reject all; Submit default; Submit personalised*).

Participants

A total of 40 participants successfully finished both experiments, with a mean age of 26.1 and standard deviation of 8.6[8] The majority, 30, had a university degree (17 Bachelor, 12 Master, 1 Doctorate). Seven had some college credit but no degree, and three a highschool diploma. 28 participants were currently studying and 12 were employed full-time. All participants were residing in the United States for the duration of the study, and did not travel to the EU. We selected this sample to prevent the confounding effects of real CMPs, which would have popped-up on top of our injected one if the participants were in the EU and thus in the regulatory scope of the GDPR. Four participants lived in the EU in the past five years, meaning they might already be familiar with pop-ups from the ePrivacy directive. All participants used Google Chrome as their main browser.

Participants were recruited through one of the author's personal network and a university mailing list. They were offered $50 upon completion of the study, and an additional $10 if they successfully recommended others.

Apparatus and Materials

The materials and apparatus of this study include a pre-study survey, a browser extension, and a post-study survey.

The pre-study survey consisted of 11 questions designed to gather demographic information (age, employment status, highest degree obtained, country of residence), check whether the participants met the study criteria (devices used to browse the web, main browser, travelling to the EU during the study), and acquire their informed consent.

To expose the participants to the different interface designs in a controlled yet ecologically realistic context, we developed a browser extension that injects different pop-ups into any website that participants would visit during their normal daily browsing (available as open-source after publication). The designs of the eight interfaces (i.e., conditions) were inspired by the designs of the top five Consent Management Platforms also used for the scraper study: QuantCast, OneTrust, TrustArc, Cookiebot, and Crownpeak.

All the text, data processing purposes, and vendor names were created by synthesising those commonly used by a random selection of those CMPS in the top 500 Alexa websites in the UK. The data processing purposes are a combination of the options that the five CMPs give to website owners when they create their own pop-up, or the purposes those websites came up with themselves. The vendor names were copied from existing websites, and picked to represent one of four categories: well-known companies (e.g, "Yahoo!"), foreign companies with English names (e.g., "Beijing Interactive Marketing"), foreign companies with non-English names (e.g., "Programatica de publicidad S.L."), and gibberish names (e.g., "s_vi_bikx7Becalgbkjkxx").

[8]Age was reported using brackets of ten years so we are unable to report the exact range; the answers were assumed to be normally distributed to calculate the mean.

The extension used the open-source JavaScript database PouchDB to store the participants' interactions with the interfaces locally, which was synchronised with a CouchDB instance running on OpenStack over an SSL encrypted connection.

The post-study survey consisted of four questions asking the participants to reflect on their general pop-up answering strategy, showed them a visualisation of their actual answers, and asked them to describe how well those answers fit their ideal preferences.

Procedure

A recruitment email was sent to potential participants asking them to join a study about web-tracker activity in the United States compared to the European Union, and answer the pre-study survey. Once approved, the participants were assigned and emailed a participant number and a link to the Chrome extension on the Chrome Web Store. After installing the extension, a welcome screen automatically appeared asking the participants to fill in their assigned number. This connected the installation to the participant number in the CouchDB database, where each participant was matched to a pre-determined experiment and condition order. Once the extension was successfully activated, a pop-up appeared notifying the participants the experiment had started. To train the participants and homogenise their understanding of the CMPs they received an additional email that informed them they might sometimes see consent pop-ups (ostensibly when they were shown the European version of a website instead of the US equivalent), explained how those pop-ups worked, and instructed them to answer the pop-ups according to their preferences.

The extension injected a pop-up every fourth url visited – including navigations on the same page, excluding automatic redirects or urls for which an answer was already recorded – to approximate the realistic frequency with which consent pop-ups are currently shown[9]. Each interface condition was repeated four times, requiring the participants to answer sixteen pop-ups per experiment. All interactions with the pop-up were recorded and timestamped: clicking on the elements, toggling purposes or vendors, scrolling the lists, navigating back and forth between the pages, submitting a consent response. Interfaces which were not interacted with were re-appended to the list of conditions and shown again for a maximum of five times, after which it was recorded as "not answered" (similar to a participant clicking or scrolling through the interface without providing a consent response). Once all conditions of the first experiment were answered, the participant progressed to the second experiment.

After completing both experiments, the participants were notified by email that the study was finished, informed that they could uninstall the extension, and asked to complete the post-study survey. The completion time of the experiment ranged from four days to three weeks, depending on how many unique urls the participant visited per day (e.g., some participants mostly visited the same websites, some went on holiday during the experiment, some installed the extension on their secondary device and only used it a couple days per week).

Data analysis

Although originally 48 participants finished the experiment, we removed eight of them because they mentioned in the survey that

[9]Based on Adzerk's Ad-Tech Insights report: [4]

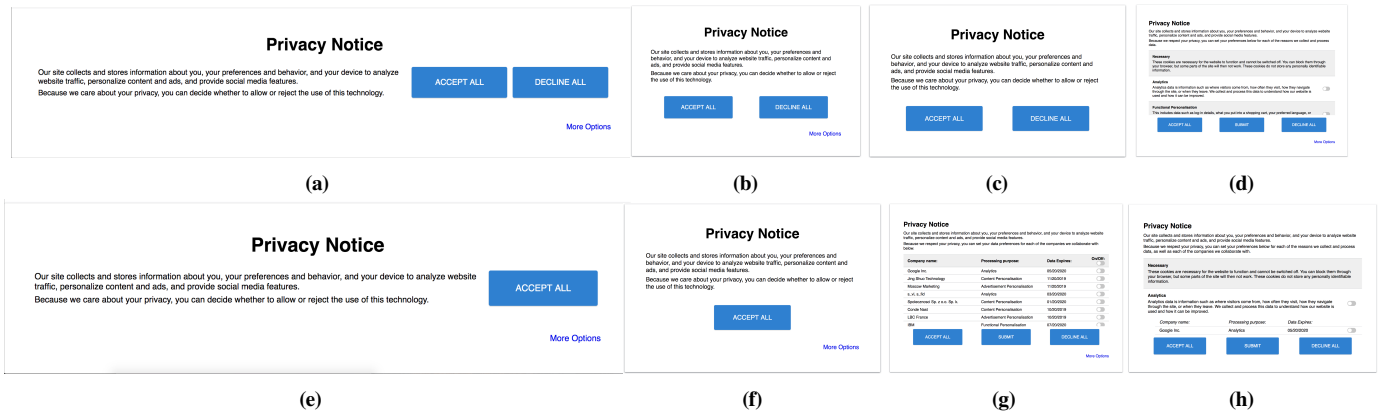


Figure 3. The 8 interface conditions: (a) Banner / Accept + Reject; (b) Barrier / Accept + Reject; (c) Bulk; (d) Bulk + Purposes; (e) Banner / Accept; (f) Barrier / Accept; (g) Bulk + Vendors; (h) Bulk + Purposes + Vendors.

their answers were affected by the study (e.g., some participants said they chose "accept all" because they wanted to give more data, despite the instructions they received). To analyse the effects of interface design on consent answers we created a linear regression model with fixed effects; we treated the participants as a factor to account for their (assumed) stable privacy preferences.

Results

General interaction patterns

Of the four possible consent choices – accept all, reject all, submit preferences, no answer – the vast majority of answers submitted by participants was through the bulk options (89.3%), with a skew towards accepting: 55.2% (707) accept all versus 34.1% (437) reject all. Just 9.7% (124) of answers represent the "submit preference" option, and 0.9% (12) were "no answer" (but recorded interactions). Of those 124 "submit" answers, merely 21 – given by 6 participants – were personalised answers, instead of submitting the default status (all toggled off). Four of those 21 were personalised by clicking the "Toggle All" button, which means *only 17 answers out of 1280 (1.3%) represent a participant consenting to a specific selection of purposes or vendors.* Whether this is because users are unable to make such decisions, are not interested in that level of detail, or fatigued by the form and frequency of the question is unclear; but it does indicate that users' consent is rarely empirically as "specific" as the GDPR requires it to be. It does not follow that specific controls should therefore be removed, but rather that such specificity could be distributed to other actors invited by the user (e.g., browser agents, consent predictors, a knowledgeable friend).

Almost all interactions (93.1%) were limited to the first page of the pop-up the participants were exposed to. Seven out of eight interfaces had a "more options" link to navigate to a second or third page for more information and granular consent choices, but this was clicked only 88 times (6.9%). When participants were exposed to a scrollable list of data collection purposes or vendors on the first or subsequent pages (560 occasions), they ignored it 68.6% (384) of the time. Of the 176 instances they did scroll, 21.6% (38) were between 0 and 25 percent of the list, and 64.2% (113) between 75 and 100 percent. In other words, anything not immediately visible to the user, anything requiring interaction to access, might as well not exist.

Notification style

The validity of one design element still under discussion by policy makers is that of the notification style [56]: a barrier in the middle of the screen which prevents the user from interacting with the website until a response is recorded, or a banner stretching the width of the screen that does not block access to the information .

We found that *notification style did not affect the consent rate of participants*. Two simple linear regressions were calculated to investigate the relationship between the answer given (accept or not) and notification style (banner or barrier). The first, comparing BARRIER to BANNER with both the Accept and Reject button, did not find a regression line at all ($F(1,279) = 0.000$, $p = 1$). The second, comparing BARRIER to BANNER with just the Accept button, found a non-significant relationship ($p = 0.702$), with a slope coefficient of 0.013 (95% CI min and max of -0.052 and 0.077 respectively) and an R^2 of 0.001.

While there was no difference in acceptance rate when participants actually answered the pop-up, the banner notification was ignored 3.6 times more often than the barrier. For this statistic we considered any pop-up that was not interacted with, but which had a time difference of at least 3 seconds between being injected and the tab being closed, as "ignored"; 133 of such instances were found, with only 21.1% (28) for the barrier and 78.9% (106) for the banner.

Button prominence

Data from our scraper indicates 'accept all' and 'reject all' buttons are not displayed with equal prominence: only a mere 12.6% of sites show both on the same page. Such unequal prominence of consent options is already considered non-compliant with the GDPR [3, 30] because it is expected they affect consent answers, but the severity of its impact is unknown.

We found that *removing the 'reject all' button from the first page increased the probability of consent by 22–23 percentage points*. We calculated two simple linear regressions to analyse the relationship between the answer given (accept or not) and the consent options on the first page (accept and reject, or just accept). The first, comparing ACCEPT ALL + REJECT ALL to ACCEPT ALL for the barrier notification, found a strong positive linear relationship between the two. The significant ($p < 0.001$) slope coefficient for the consent answer was 0.220, meaning the accept rate increased

on average by 22.0 percentage points when the reject all button was removed from the first page. The 95% CI had a minimum and maximum of 0.149 and 0.290 respectively. The R^2 was 0.117, so 11.7% of the variation in answers for the barrier notification can be explained by the changing prominence of the buttons.

The second regression compared ACCEPT ALL + REJECT ALL to ACCEPT ALL for banner notifications and found a similarly strong, positive linear relationship between the button prominence and answer given. The significant ($p < 0.001$) slope coefficient was 0.231, meaning the accept rate increased on average by 23.1 percentage points when the reject all button was removed from the first page. The 95% CI had a minimum and maximum of 0.163 and 0.230 respectively. The R^2 was 0.135, so 13.5% of the variation in answers for the banner notification can be explained by the changing prominence of the buttons.

Level of granularity

The most common order in which consent options are displayed is bulk first, followed by the data collection purposes on the second page and the vendors on the third page, or some combination of those two on the same page.

We found that *displaying more granular consent choices on the first page decreased the probability of consent by 8–20 percentage points*. We calculated a simple linear regression to compare a BULK only interface to an interface that combined BULK + PURPOSES; BULK + VENDORS; and BULK + PURPOSES + VENDORS on the same page. We found a significant ($p < 0.01$) negative relationship between all increases in the level of granularity of consent options and the answer given, with different strengths depending on the kind of options that were available. As illustrated by Table 2, showing just the vendors affected the acceptance rate the most (−0.200), whereas just the purposes (−0.088) and the combination of vendors and purposes (−0.119) were closer together but still lower than the baseline interface with just bulk options. Along the same lines, the 95% CIs overlap most between PURPOSES and PURPOSES + VENDORS and only a little with VENDORS.

Table 2. Level of granularity on the first page, with bulk consent as the reference

	Dependent variable:	*95% CI:*
	'accept all' clicked	lower : upper
Bulk + Purposes	−0.088**	−0.151 : −0.024
	(0.032)	
Bulk + Vendors	−0.200***	−0.263 : −0.137
	(0.032)	
Bulk + Purposes + Vendors	−0.119***	−0.182 : −0.056
	(0.032)	
Observations	640	
R^2	0.062	
F Statistic	13.210*** (df = 3; 597)	
Note:		*p<0.1; **p<0.01; ***p<0.001

Participant Strategies and Behaviour Patterns

While the experimental data suggests how different designs affect how "freely given" the consent answers of participants are, it does not provide information about how those answer relate to their preferred privacy settings. In a post-study survey, we requested participants to describe their overall answering strategy, showed them a visualisation of their actual behaviour, and then asked them to state how well their answers reflected their ideal preferences and, if not, why. To structure these findings, we classify participants according to their general consent answers: always accept, mostly accept (>= 75%), mixed consent, mostly reject (>= 75%), always reject.

When asked what they based their choices on, the answers touched on eleven different topics. The four 'always accept' participants cited a general apathy towards privacy concerns and "*just did it to make the window go away*". The one participant that 'always rejected', no matter whether that required more effort, argued that they would only accept data collection if it was to use a particular feature offered by the site. The eleven participants categorised as 'mostly reject' heavily emphasised a disagreement with the practice of tracking in general and stated they would only consent to have their data collected if it was for websites they trusted. Two of those also mentioned that they did not feel a need for any personalisation. The participants that fell into the 'mostly accept' and 'mixed consent' category were more diverse. Most often mentioned were pragmatic reasons such as just wanting to get to the site as quickly as possible, not believing the controls were meaningful, and not wanting to lose any functionality. Eight decided based on trust, whether it was the website or the vendors, and the sensitivity of the data they would be submitting (e.g., banking information). One participant stated that they relied on other methods to protect their privacy, so did not care that much about their pop-up answers: "*I tend to vary my devices/browsers/accounts/use incognito and duckduckgo a lot, I'm not too worried about my data being tracked to every detail*."

After being shown a visualisation of their actual consent behaviour and asked if it matched their ideal settings, the responses were predominately that it did not. Only those falling into the two extreme categories – 'always accept' and 'always reject' – all indicated they agreed (3) or strongly agreed (2) with their answers. For the remaining three categories, the sentiments were mostly spread evenly along the spectrum, with 11 somewhat agreeing, 3 neither agreeing nor disagreeing, 7 somewhat disagreeing, 1 disagreeing, and 3 strongly disagreeing.

The 25 participants who indicated their behaviour did not match their ideal privacy settings were asked to explain what the reason for this difference was. Participants mentioned desires such as just wanting more privacy ("*I would rather companies not collect any information*"); the fear of unknown consequences of opting-out ("*I didn't want to risk the website not working after that*"); and not knowing what their ideal preferences even are. The most common reason mentioned, however, was the interface design. Participants lamented the fact that pop-ups stand in the way of their primary goal (accessing a service), that the frequency of the pop-ups caused frustration and consent fatigue, and even the perception that the pop-up "*forced them to accept*" – even though these options were available on the second page.

Interim Discussion

The experimental results indicate how two of the most common consent interface designs – not showing a 'reject all' button on the first page; and showing bulk options before showing granular

control – make it more likely for users to provide consent, violating the principle of "freely given"[10]. The notification style, on the other hand, appears to have no effect on the answer, but possibly a large effect on whether an answer is given at all, suggesting that a non-blocking mechanism provides a desired third consent option to users: a neutral middle-ground. The qualitative reflections of the participants, however, put into question the entire notice-and-consent model not because of specific design decisions but merely because an action is required before the user can accomplish their main task and because they appear too frequently if they are shown on a website-by-website basis.

Limitations

The participant sample is by no means representative of the general population in the United States: they are almost all young and university-educated, and recruited primarily through an emailing list of a computer science department. Arguably, this means that our results describe a "best case scenario": these participants should be more knowledgeable about privacy issues and better equipped to understand consent interfaces than the average web user.

There are a number of confounding variables that could have affected the participants' answers. First, although the condition order was counterbalanced, we cannot guarantee that the participants were actually exposed to them in that order (e.g., if they opened multiple tabs in a row and visited them anachronistically), meaning order effects might not be controlled for. Second, because we showed the same pop-up to each participant until we recorded four answers per interface, some participants were exposed to the different conditions more often than others. Lastly, participants might have also encountered "real" pop-ups at the same time as the injected ones if the website they were visiting was within the territorial scope of the GDPR.

While the GDPR is a European policy, our experiments were conducted in the United States. These populations have been exposed to different legal regimes and different consent controls over the year, something which we expect has affected their mental model of these kind of pop-ups and accordingly, how they answer them. This might influence the extent to which these findings can be generalised to a European population, and thus how they should be used to inform EU policy changes.

DISCUSSION AND CONCLUSION

The results of our empirical survey of CMPs today illustrates the extent to which illegal practices prevail, with vendors of CMPs turning a blind eye to — or worse, incentivising —- clearly illegal configurations of their systems. Enforcement in this area is sorely lacking. Data protection authorities should make use of automated tools like the one we have designed to expedite discovery and enforcement. Designers might help here to design tools for regulators, rather than just for users or for websites. Regulators should also work further upstream and consider placing requirements on the vendors of CMPs to only allow compliant designs to be placed on the market. Such enforcement may be possible as the Court of Justice indicates that plugin system designers can be 'joint controllers' along with websites [17, 37, 54], and the UK's ICO indicates it may be willing to force advertising trade bodies to alter their standards [31]. If this is the case, regulators must carefully consider how to build a robust and well-maintained evidence base for user-centric CMP design.

A core takeaway from the user study is that placing controls or information below the first layer renders it effectively ignored. This leaves a few options for genuine control of tracking online. If the notice-and-consent model is to continue, it may be necessary to declare that, for example, consent can never be valid with the presence of the (on average) hundreds of third parties we have shown data is sent to and cookies laid by today. This would mean consent would only be valid if a compact but representative and rich description can be placed on the first layer, and could certainly be a possible direction for the Court of Justice to consider if they interpret the principles of data protection in a future case.

An alternative approach would be to overhaul the design pattern of the consent banner or barrier, and have richer, more durable ways to set preferences, potentially within the browser. The key is that such browser settings would be legally binding, rather than weak and self-regulatory in nature. Yet the current heavy lobbying around the EU's draft ePrivacy Regulation has centred in part on adtech firms trying to prevent browser settings having legally binding effect — part of an ongoing drama for many years about the potential legal status of 'Do Not Track' signals [33]. Designers have a role here: how can users reflect on tracking across the Web, rather than on a per-site basis? If users are not to automatically reject everything, how can advertisers negotiate and present them with reasons that they should consent? Might there be a role for delegation of preferences to a trusted civil society actor, and what kind of relationship, information and interaction might the user have with these? We invite and encourage researchers to bring their skills and views to bear on these important, current issues at the confluence of regulation, design and fundamental rights.

[10]It should be noted this data alone is not enough to establish legal compliance.

ACKNOWLEDGMENTS

Our thanks to José Juan Dominguez Veiga and Kristian Borup Antonsen for their invaluable assistance programming the scraper and extension, and to Luke Taylor, Ignacio Avellino, and Benjamin Cowan for their advice with the statistical analysis.

Midas Nouwens was supported by the Aarhus Universitets Forskningsfond and the Danish Agency for Science and Higher Education. Ilaria Liccardi was supported by the William and Flora Hewlett Foundation and the NSF 1639994: Transparency Bridges grant. Michael Veale was supported by the Alan Turing Institute under EPSRC grant no. EP/N510129/1.

REFERENCES

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang, and Shomir Wilson. 2017. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *ACM Comput. Surv.* 50, 3, Article 44 (Aug. 2017), 41 pages. DOI: http://dx.doi.org/10.1145/3054926

[2] Alessandro Acquisti and Jens Grossklags. 2005. Privacy and rationality in individual decision making. *Security Privacy, IEEE* 3, 1 (2005), 26–33.

[3] Advocate General Szupunar. 2019. Case C-673/17 *Planet49 GmbH v Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband e.V.* ECLI:EU:C:2019:246, Opinion of the Advocate General. (2019).

[4] Adzerk. 2019. Adtech Insights — August 2019 Report. (2019). https://adzerk.com/assets/reports/AdTechInsights_Aug2019.pdf

[5] Julio Angulo, Simone Fischer-Hübner, , Tobias Pulls, and Erik Wästlund. 2011. Towards Usable Privacy Policy Display & Management for PrimeLife. *S. M. Furnell, & N. L. Clarke (Eds.), Proceedings of international symposium on human aspects of information security & assurance (HAISA 2011)* (2011), 108 – 117.

[6] Article 29 Working Party. 2018. *Guidelines on Consent under Regulation 2016/679 (WP259 rev.01)*. European Union.

[7] Autoriteit Persoonsgegevens. 2019. *Hoe Legt de AP de Juridische Normen Rond Cookiewalls Uit?* AP, Den Haag.

[8] Meinert David B., Dane K. Peterson, John R. Criswell, and Martin D. Crossland. 2006. Towards Usable Privacy Policy Display & Management for PrimeLife. *Journal of Electronic Commerce in Organizations (JECO)* 4, 1 (2006), 1–17.

[9] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. 2016. Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proceedings on Privacy Enhancing Technologies* 2016, 4 (2016), 237–254.

[10] Axel Bruns. 2019. After the 'APIcalypse': Social Media Platforms and Their Fight against Critical Scholarly Research. *Information, Communication & Society* 22, 11 (2019), 1544–1566. DOI: http://dx.doi.org/10.1080/1369118X.2019.1637447

[11] Tania Bucher. 2013. Objects of Intense Feeling: The Case of the Twitter API : Computational Culture. *Computational Culture: A Journal of Software Studies* 3 (2013). http://computationalculture.net/objects-of-intense-feeling-the-case-of-the-twitter-api/

[12] Fred H Cate. 2010. The limits of notice and choice. *IEEE Security & Privacy* 8, 2 (2010), 59–62.

[13] Damian Clifford, Inge Graef, and Peggy Valcke. 2019. Pre-formulated Declarations of Data Subject Consent—Citizen-Consumer Empowerment and the Alignment of Data, Consumer and Competition Law Protections. *German Law Journal* 20, 5 (2019), 679–721.

[14] Commission nationale de l'informatique et des libertés (CNIL). 2019. Délibération n° 2019-093 du 4 juillet 2019 portant adoption de lignes directrices relatives à l'application de l'article 82 de la loi du 6 janvier 1978 modifiée aux opérations de lecture ou écriture dans le terminal d'un utilisateur (notamment aux cookies et autres traceurs) (rectificatif). (2019).

[15] Gregory Conti and Edward Sobiesk. 2010. Malicious Interface Design: Exploiting the User. In *Proceedings of the 19th International Conference on World Wide Web*. ACM, 271–280.

[16] Jake R. Conway, Alexander Lex, and Nils Gehlenborg. 2017. UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties. *Bioinformatics* 33, 18 (2017), 2938–2940. DOI: http://dx.doi.org/10.1093/bioinformatics/btx364

[17] Court of Justice of the European Union. 2019a. Case C-49/17 *Fashion ID GmbH & Co.KG v Verbraucherzentrale NRW eV.* ECLI:EU:C:2019:629. (2019).

[18] Court of Justice of the European Union. 2019b. Case C-673/17 *Planet49 GmbH v Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband e.V.* ECLI:EU:C:2019:801. (2019).

[19] Lorrie Cranor. 2002. *Web privacy with P3P*. O'Reilly Media, Sebastopol, CA.

[20] Lorrie Faith Cranor. 2012. Necessary but Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice The Economics of Privacy. *Journal on Telecommunications and High Technology Law* 10, 2 (2012), 273–308.

[21] Mark R. Warner Deb Fisher. 2019. Deceptive Experiences To Online Users Reduction (DETOUR) Act. https://www.scribd.com/document/405606873/Detour-Act-Final

[22] Martin Degeling, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz. 2018. We Value Your Privacy... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. *arXiv preprint arXiv:1808.05096* (2018).

[23] European Data Protection Supervisor. *EDPS Opinion on the Proposal for a Regulation on Privacy and Electronic Communications (ePrivacy Regulation), Opinion 6/2017*. EDPS, Brussels, BE.

[24] European Union. 1995. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 1995 L 281/31. (1995).

[25] European Union. 2002. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) OJ L 201. (2002).

[26] European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1. (2016).

[27] Brian J Fogg. 2009. A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology*. ACM, 40.

[28] Forbrukerrådet. 2019. Deceived by Design: How tech companies use dark patterns to discourage us from exercising our rights to privacy. (2019). https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf

[29] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 534.

[30] Information Commissioner's Office. 2019a. *Guidance on the Use of Cookies and Similar Technologies*. ICO, Wilmslow, Cheshire.

[31] Information Commissioner's Office. 2019b. *Update Report into Adtech and Real Time Bidding*. ICO, Wilmslow, Cheshire.

[32] Carlos Jensen and Colin Potts. 2004. Privacy policies as decision-making tools: an evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 471–478.

[33] Irene Kamara and Eleni Kosta. 2016. Do Not Track Initiatives: Regaining the Lost User Control. *International Data Privacy Law* 6, 4 (2016), 276–290. DOI: http://dx.doi.org/10/gdxwds

[34] 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*. ACM, 4.

[35] Eleni Kosta. 2013. Peeking into the Cookie Jar: The European Approach towards the Regulation of Cookies. *International Journal of Law and Information Technology* 21, 4 (2013), 380–406. DOI: http://dx.doi.org/10.1093/ijlit/eat011

[36] A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister. 2014. UpSet: Visualization of Intersecting Sets. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 1983–1992. DOI: http://dx.doi.org/10.1109/TVCG.2014.2346248

[37] Rene Mahieu, Joris van Hoboken, and Hadi Asghari. 2019. Responsibility for Data Protection in a Networked World: On the Question of the Controller, Effective and Complete Protection and Its Application to Data Access Rights in Europe. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 10, 1 (2019), 84–104.

[38] Arunesh Mathur, Gunes Acar, Michael J Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 81.

[39] Célestin Matte, Nataliia Bielova, and Cristiana Santos. 2019. Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework *(Under submission)*. https://arxiv.org/abs/1911.09964v1

[40] John McCarthy. 2019. Over 90% of users consent to GDPR requests says Quantcast after enabling 1bn of them. https://www.thedrum.com/news/2018/07/31/over-90-users-consent-gdpr-requests-says-quantcast-after-enabling-1bn-them. (2019).

[41] A. M. McDonald and L. F. Cranor. 2008. The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society* 4 (2008), 540 – 565.

[42] H. Nissenbaum. 2011. A contextual approach to privacy online. *Daedalus* 140, 4 (2011), 32–48.

[43] Jonathan A. Obar and Anne Oeldorf-Hirsch. 2018. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society* 0, 0 (2018), 1–20. DOI:http://dx.doi.org/10.1080/1369118X.2018.1486870

[44] Robert W Reeder, Lujo Bauer, Lorrie Faith Cranor, Michael K Reiter, Kelli Bacon, Keisha How, and Heather Strong. 2008. Expandable grids for visualizing and authoring computer security policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1473–1482.

[45] Iskander Sanchez-Rola, Matteo Dell'Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos. 2019. Can I Opt Out Yet?: GDPR and the Global Illusion of Cookie Control. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security (Asia CCS '19)*. ACM, New York, NY, USA, 340–351. DOI: http://dx.doi.org/10.1145/3321705.3329806

[46] 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.

[47] Natasha Singer. 2016. When Websites Won't Take No for an Answer. *New York Times* (15 5 2016). Retrieved Sept 19, 2019 from https://www.nytimes.com/2016/05/15/technology/personaltech/when-websites-wont-take-no-for-an-answer.html?mcubz=0&_r=0

[48] Jannick Sørensen and Sokol Kosta. 2019. Before and After GDPR: The Changes in Third Party Presence at Public and Private European Websites. In *The World Wide Web Conference (WWW '19)*. ACM, New York, NY, USA, 1590–1600. DOI: http://dx.doi.org/10.1145/3308558.3313524

[49] European Data Protection Supervisor. 2018. EDPS Opinion on the legislative package "A New Deal for Consumers". https://edps.europa.eu/sites/edp/files/publication/18-10-05_opinion_consumer_law_en.pdf

[50] Richard H Thaler and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

[51] Oisin Tobin. 2019. Cookie consent revisited. *Privacy and Data Protection* 19 (2019), 11. Issue 5.

[52] Martino Trevisan, Stefano Traverso, Eleonora Bassi, and Marco Mellia. 2019. 4 Years of EU Cookie Law: Results and Lessons Learned. *Proceedings on Privacy Enhancing Technologies* 2019, 2 (2019), 126–145.

[53] 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 (CCS '19)*. ACM, New York, NY, USA, 973–990. DOI: http://dx.doi.org/10.1145/3319535.3354212

[54] Brendan Van Alsenoy. 2019. *Data Protection Law in the EU: Roles, Responsibilities and Liability*. Intersentia, Cambridge.

[55] Tony Vila, Rachel Greenstadt, and David Molnar. 2003. Why We Can'T Be Bothered to Read Privacy Policies Models of Privacy Economics As a Lemons Market. In *Proceedings of the 5th International Conference on Electronic Commerce (ICEC '03)*. 403–407.

[56] Frederik J Zuiderveen Borgesius, Sanne Kruikemeier, Sophie C Boerman, and Natali Helberger. 2017. Tracking Walls, Take-It-Or-Leave-It Choices, the GDPR, and the ePrivacy Regulation. *European Data Protection Law Review* 3, 3 (2017), 353–368. DOI: http://dx.doi.org/10/gfsh4x