I Refuse if You Let Me: Studying User Behavior with Privacy Banners at Scale

I Refuse if You Let Me: Studying User Behavior with Privacy Banners at Scale

Nikhil Jha[†], Martino Trevisan[‡], Marco Mellia[†], Rodrigo Irarrazaval[⋆], Daniel Fernandez[⋆]
[†]Politecnico di Torino, [‡]University of Trieste, [⋆]Illow
first.last@{polito,dia.units}.it, {daniel,rodrigo}@illow.io

***Abstract*—Privacy Banners are a common experience while surfing the Web. Mandated by privacy regulations, they are the way for users to express their consent to the usage of cookies and data collection. They take various forms, carry different wordings and offer different interaction mechanisms. While several works have qualitatively evaluated the effectiveness of privacy banners, it is still unclear how users take advantage of the options offered and if and how the design of the banner could influence their choice.**

This work presents a large-scale analysis of how the Privacy Banner options impact on users' interaction with it. We use data from a global Consent Management Platform serving more than 400 websites with visitors from all countries. With this, we observe more than 4 M interactions collected over three months. We find that only 1-4% of visitors opt out of cookies when more than one click is required. Conversely, when offered a `Reject All` button to deny consent with a single click, the percentage of users who deny consent increases to about 21%.

We further investigate other properties, such as the visitor's country, device type, banner position, etc. While the results confirm some common beliefs, to the best of our knowledge, this is the first work to accurately quantify how people interact with Privacy Banners and observe the effect of offering a single-click refusal option. We believe our work improves the understanding of user behaviour and perception of privacy, as well as the implications and effectiveness of privacy regulations.

***Index Terms*—Cookie, Privacy, Consent Management Platforms, Web Measurements**

I. Introduction

In the Web ecosystem, most services monetize the content they offer via online advertising. This has led to a massive, unprecedented collection of personal data, which is essential for behavioural or targeted advertising and for marketing and business analytics. This scenario created tension between the online industry and users around their privacy.

The collection of personal information often relies on the use of cookies. Cookies are pieces of text stored in a client's browser set by the visited website. By retrieving previously set cookies, a website can recognize the user and improve one's experience, e.g., by remembering the language or the preferred theme. However, cookies (and more advanced mechanisms [1, 2]) are also used to collect information about users, track them across different websites, and leverage the information for not only personalized ads [3–7]. This threatens users' privacy, and the research community has started proposing alternatives [8].

On their hand, public bodies and regulators have started proposing and enforcing regulations to govern the phenomenon. The European Union (EU) was the first to enact a privacy law that applies to a large geographic region. With the 2009 "Cookie Law" directive [9] all websites that use first-party or third-party cookies to track users' behaviour must obtain user consent via a *Privacy Banner* – and must not use cookies the user has refused. New requirements and obligations have been added with the adoption of the General Data Protection Regulation (GDPR) in 2018 [10]. At the time of writing, 137 out of 194 countries had put in place legislation to secure the protection of data and privacy[1]. To simplify the deployment of Privacy Banners, new companies offer web administrators simple technical solutions called Consent Management Platforms (CMPs) [11] to ease compliance with privacy regulations.

Although regulations state that users shall *freely* provide their consent, recent works have claimed that the way the Privacy Banners present options impacts users' choices. Unsurprisingly, most Privacy Banners encourage users to provide consent [12, 13], i.e., making it simpler to accept than to refuse. Given the novelty of the problem and the lack of data, few studies have focused on these types of interactions, and most provided qualitative evidence or small-scale experiments.

In this paper, for the first time to the best of our knowledge, we present a large-scale study of how users interact with Privacy Banners. We leverage data from a global, medium-size CMP present on hundreds of websites visited by users worldwide. By analyzing more than 4 million interactions, we observe the factors influencing users' decision to accept or reject.

We first find that the options the Privacy Banner offers play a fundamental role. When forced to go through some customization window to deny their consent, the percentage of users who do so is about around 1–4% depending on the region they are connecting from. That is, visitors simply select the `Accept All` option to remove the banner from the screen quickly. Conversely, when offered the option to deny their consent with a single click, i.e., with a `Reject All` button, the percentage of users doing so suddenly grows to about 21%. This has clear implications for the Internet economy, which bases its revenue on the ability to collect data.

We also investigate other aspects that may impact users' choices. For instance, we unexpectedly find that Apple iOS

978-3-903176-58-4 ©2023 IFIP

[1]https://unctad.org/page/data-protection-and-privacy-legislation-worldwide – visited in March 2023

users are more likely to accept than Android users. Moreover, only a handful of visitors do actually check the privacy or cookie policy text: It questions the effectiveness of the Privacy Banner as an instrument for collecting informed consent from users.

The remainder of the paper is organized as follows. Section II summarizes related work, while Section III describes the dataset we use and the processing steps we design to avoid bias in the measurements. We next show and discuss our results in Section IV, while Section V states the limitations of our work and concludes the paper.

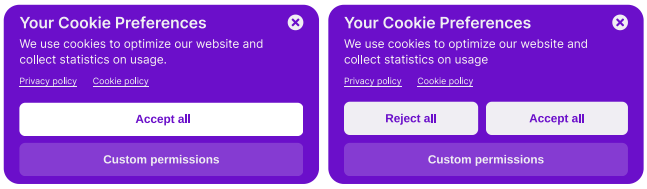
II. Related Work

The impact of privacy regulations on the web ecosystem has been studied from various points of view. The introduction of the Cookie Law [9] in 2013 led to the proliferation of Privacy Banners [14]. When a user visits a website for the first time, they have to interact with the Privacy Banner, and, only after getting the explicit user's consent, the website (and any third party embedded in the website) could install cookies and start the data collection. Privacy Banners, however, do not fully protect users in many cases [15]. Later in 2018, the GDPR has profoundly influenced the Internet user experience [16–20], at least for EU-based users, also defining severe sanctions for violators. Most websites base their business model on advertising, which, in turn, requires that users accept the use of cookies and the collection of personal data. Thus, some websites and CMPs make efforts to increase the cookie acceptance rate. Recent research has shown that banners often nudge users to acceptance by exploiting dark patterns in the user interface, if not openly disregarding GDPR's requirements [12], or making it difficult for users to exercise their rights [21]. They also hinder automated web measurement, hiding the true content of a website, which is visible only upon cookie acceptance [22].

Nudging includes offering the user a `Accept All` default button via intrusive banners [23, 24], which is often the case [25] with websites presenting large pop-ups or wall-style banners that cover most of the webpage content. Researches have shown that apparently minor design choices have a significant effect on inducing the user to accept the cookies [13, 26–30].

In general, it has been shown that most users tend to ignore privacy-related notices [31–33], up to getting annoyed by these. This behaviour has gone under the name of "privacy paradox": Users claim to be concerned about their privacy, while at the same time taking little actions to protect their data [34].

Our work is complementary to this body of literature. Previous works provide only qualitative results or small-scale measurements to support their claims. In this work, we have the possibility of exploiting data from a medium-sized CMP. We are among the first to directly measure at scale how users actually interact with consent banners, confirming some common belief, but precisely quantifying it with thorough measurements.



(a) Default banner without the `Reject All` button.

(b) Banner with the `Reject All` button.

Fig. 1: Privacy Banners presented to users.

III. Measurements and Dataset

A. The Consent Management Platform

In this paper, we rely on data collected within a medium-sized CMP. It provides web developer with the ability to install a simple Privacy Banner to enable/disable data collection via cookies or other advanced means. The banner takes the form of a small overlay window that can be placed in different parts of the screen. We show it in Figure 1. The shape is the same on both desktop and mobile devices. The user is offered an `Accept All` button to accept all cookies at once and a `Custom Permissions` button (Figure 1a). This brings the user to a second window where they can select which cookies to accept from a short list of categories. These include (i) necessary , (ii) statistical, (iii) preferential, and (iv) marketing cookies. Necessary cookies cannot be deactivated as they are vital for the website operation. Depending on the website, the Privacy Banner is shown on the top or on the bottom of the webpage. In the latter case, the website administrator can choose to show it as a rectangle (default behavior, as in Figure 1a) or in a square shape in the bottom-left corner of the screen. At last, the banner offers direct links to the website cookie and privacy policy. Both policies contain details about which data the site collects and for what purposes, and which cookies the system uses, including third-party ones.

The `Reject All` button: The latest practices regarding cookie management in GDPR countries recommend the Privacy Banners to offer a `Reject All` button. This is a consequence of the fine imposed by CNIL (the French data protection authority) on Google and Facebook in January 2022[2]. The two companies were fined for using confusing language in their Privacy Banners, and for making it difficult to opt out of cookie usage. In fact, it was not as easy to reject cookies as it was to accept them, and this was considered a form of dark pattern that nudges users to provide their consent. Since the last week of August 2022, the CMP analyzed in this study has updated its solution to offer a `Reject All` button (Figure 1b). If selected, the system will disable all cookies except the necessary ones. The button bears the text `Reject All` and has a similar shape and style as the

[2]https://www.cnil.fr/en/cookies-cnil-fines-google-total-150-million-euros-and-facebook-60-million-euros-non-compliance – visited in March 2023

`Custom Permissions` button. This button is only shown to visitors of GDPR countries after August 25, 2022.[3]

B. CMP's Data Collection

The CMP collects data when users interact with the Privacy Banner shown on a website. The collection happens when the user submits their preference. No data is collected if users do not perform any action on the banner. In details, after the user's selection, the CMP sets a necessary cookie on user's browser to store their preference, and logs data about the time of the visit, the website showing the banner, which cookies the user accepted, the user agent offered by the browser, and the user country of origin, obtained through the client IP /24 subnet geolocation via the MaxMind GeoIP[4] database.[5] This information is necessary to implement the functionalities of the platform (i.e., record user's preferences for the next visits to the website), and it is useful to customize the information provided to users (e.g., show the banner in different languages, show the `Reject All` button if needed), to collect statistics about the usage of the platform, and to bill the website deploying the CMP. All these pieces of information are documented in the privacy policy the CMP offers to users.

Each user who submits (or changes) their preferences generates an entry in the log, which we call *interaction*. Each entry is associated with a random user-id. This makes it impossible to re-identify or track a user across different websites, guaranteeing user's privacy. To further protect the customers of the CMP, the website name is also anonymized by replacing the domain name with a random identifier.

Ethical Aspects: In this study, we adopted a lawful and ethical methodology for data collection and processing. First of all, users who interact with the Privacy Banner must accept technical cookies and thus the privacy policy. Indeed, technical cookies are mandatory to store the user's choice. As said, the Privacy Banners explicitly list "carrying out statistics, managing incidents or conducting market studies" as one of the data collection purposes. Our work fits this purpose. Conversely, we do not collect any data for those users who did not accept technical cookies and thus the privacy policy. Second, we argue that the data we collect can hardly be considered "personal data". We only collect the $/24$ subnet and the user agent to extract user's country and device. Neither the $/24$ subnet nor the user agent are personal data and do not carry information relating to an identified or identifiable natural person.

C. Data Collection and Pre-Processing

We conduct our analysis from the 1st of July 2022 to the beginning of October 2022. In total, we observe 4 million interactions generated by users that interacted with the CMP banner at least once on the 434 websites recorded during the three-month measurement period. Most visitors (93%) are located in South America (where the main business of the CMP is). The remaining ones come from other continents, and we breakdown the audience provenience in Table II. We consider and properly address this unbalance in the data for our upcoming analysis to provide results which are not biased by the unequal distribution of countries.

TABLE I: Summary of the two periods we use to compare user behavior on Privacy Banners with or without the `Reject All` button.

Period	Start	End	`Reject All` button
Period A	Jul 1, 2022	Aug 24, 2022	Not present
Period B	Aug 25, 2022	Oct 4, 2022	Only for users from GDPR countries

TABLE II: Number of interactions per geographical region.

Region	# of interactions	% of interactions
Latin America	3 750 135	93.28%
North America	153 365	3.81%
GDPR-regulated	71 640	1.78%
Africa	31 917	0.79%
Asia	7 722	0.19%
Oceania	2 782	0.07%
Rest of Europe	2 691	0.07%

Websites belong to different categories, including e-commerce portals and educational institutions. Globally, the CMP manages between 20 k to 30 k new interactions on a daily basis – i.e., new users that come across the CMP Privacy Banner and interact with it.

For each interaction, we compute the choice the user performed according to the combination of accepted cookie categories. In details, we classify interactions as:

- *Accepted-All*: if all cookie categories were accepted, either with a single click on the `Accept All` button, or by individually accepting all the cookies after clicking on `Custom Permissions` button;
- *Mandatory-Only*: if only the necessary cookies were accepted, either by clicking `Reject All` button if present, or by manually deactivating all the cookies after clicking on `Custom Permissions` (with the exception of necessary cookies);
- *Custom*: if at least one among the optional statistical, preferential and marketing cookies was accepted through the `Custom Permissions` screen.

For simplicity, we introduce the class *Reject-Some* to indicate the union of *Mandatory-Only* and *Custom*. These include all interactions but *Accepted-All* – i.e., those in which the user did not accepted all cookies.

To analyze the impact of the presence of the `Reject All` button, we define two measurement periods as detailed in Table I. The first period extends from the beginning of July to August 24, 2022. During this period, the Privacy Banner only

[3]"GDPR countries" refers to any European country where the GDPR is in place. This includes U.K. which adopted GDPR in the "Data Protection Act" in 2018.

[4]https://www.maxmind.com/

[5]We do not consider IPv6, as it generates negligible traffic.

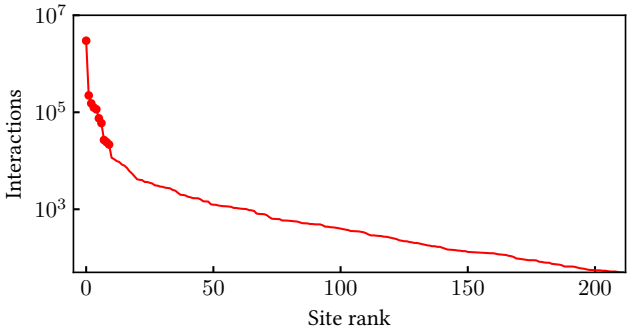


Fig. 2: The number of interactions for each website. The markers indicate the values for the top-10 websites.

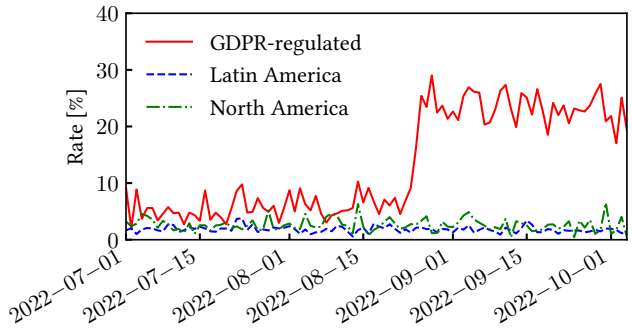


Fig. 3: Temporal evolution of the per-region *Reject-Some* rate.

included the `Accept All` and `Custom Permissions` buttons as shown in Figure 1a. We call it *Period A*. The second period starts on August 25, 2022 and ends on October 4, 2022. Here, visitors from any GDPR-regulated countries face the new version of the Privacy Banner with the additional `Reject All` button, sketched in Figure 1b. We refer to this period as *Period B*. We use these two periods to contrast user's behavior with different options in the Privacy Banner. In particular, our dataset allows us to measure the extent to which users reject cookies when the banner provides an immediate opportunity to do so (or not).

D. Dataset Analysis

We now briefly describe the dataset and detail the analysis methodology we design to avoid possible bias in the study. The CMP is present on 434 websites that have a very different audience. Some of them are very popular and generate more than 1 M interactions in total. To characterize the website popularity, we show the volumes of interactions per website in Figure 2. Sites are sorted in decreasing number of interactions (notice the log scale on the y-axis). We observe that top websites receive most of the interactions. We record 222 websites collecting less than 50 interactions.

Given the large imbalance in the website audience, we want to prevent large websites from biasing the results. For this, we opt to show results using a website-wise macro-average of the metrics under study. In other words, we compute the desired metric separately for each website. Then we compute the average over the websites. In such way, each website has the same weight in the final metric, regardless of the number of interactions it received.

Formally, given a target metric M, a set of websites $\mathcal{W}$, a population of interactions on a website $\mathcal{I}_w$, a function $\mathcal{M}(M, i)$ which return 1 if i refers to M, 0 otherwise (e.g., whether interaction i records a *Reject-Some* choice or not), we define as $\overline{M}(\mathcal{I})$ the website-wise macro-average of M computed over the samples belonging to the subset $\mathcal{I} = \bigcup_{w \in W} \mathcal{I}_w$:

$$\overline{M}(\mathcal{I}) = \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} \left[\frac{1}{|\mathcal{I}_w|} \sum_{i \in \mathcal{I}_w} (\mathcal{M}(M, i)) \right]. \quad (1)$$

Together with the macro-average, we also evaluate a confidence interval of such average. Hence, each estimate is presented as:

$$\overline{M}(\mathcal{I}) \pm c \cdot \frac{\overline{S}(\mathcal{I})}{|W|},$$

where c corresponds to the quantile of a Student's t-distribution with $|W| - 1$ degrees of freedom, and $\overline{S}(\mathcal{I})$ is the sample standard deviation of each website-wise average. In this work, we consider a confidence interval of 90% and report the confidence interval as an error bar. As our main target metric we consider the *Reject-Some* rate.

IV. Results

In this section, we present our results. We first dissect user behaviour by geographic region and show the impact of adding the `Reject All` button in GDPR countries. Next, we investigate the role of other factors, such as user device and privacy banner position. Finally, we examine the behaviour of users who have particular interactions with the Privacy Banners, i.e., custom choices (*Custom* interactions) or access to the website privacy policy.

A. Region-wise temporal analysis

We first show the evolution of the *Reject-Some* rate over time in Figure 3, separately for the three most represented geographic regions in our dataset. Here, for each day, we compute the *Reject-Some* rate for each website (and region) and then average the values to obtain the macro average. Notice that it sums both the *Mandatory-Only* and *Custom* rates. To avoid websites with very few interactions affecting the results, we evaluated the per-day average only on the websites recording at least 10 interactions on that day. We first observe that the rate exhibits a flat trend for North and Latin America and settles to values in the order of 2%. In European countries where GDPR is in force (solid red line),

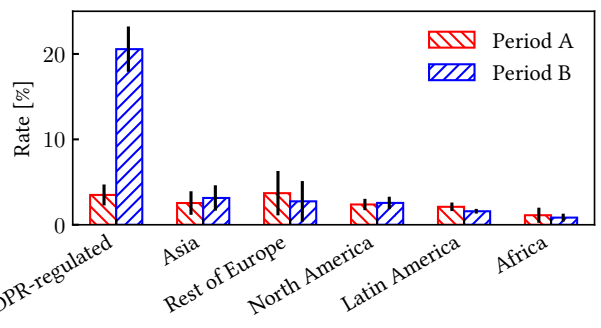


Fig. 4: Users' *Reject-Some* rate in different continents, before and after the introduction of `Reject All` button in GDPR countries.

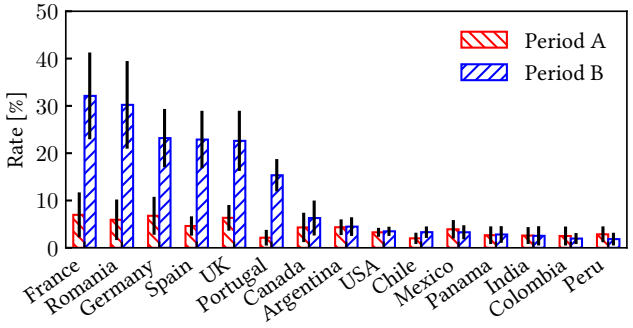


Fig. 5: *Reject-Some* rates according to user's country, sorted in descending order by the rate in *Period B*.

the *Reject-Some* rate is in the order of 3.5% until August 24 and then jumps over 20%. This increase corresponds to the transition between *Period A* and *Period B* and provides a first quantification of the impact of the `Reject All` button. In the following, we will analyze this in depth.

Notice that new websites have become CMP customers during the observation period (while few have left the CMP). The trend has been increasing over the months as many new websites have been more numerous than desertions. While on the first weeks of July 2022, we find approximately 50 websites every day with more than 50 interactions, on the first week of October 2022, this number increases to $\approx$ 120. Finally, we observe that the *Accepted-All*, *Custom* and *Mandatory-Only* rates do not depend on the website popularity. If we compute a linear regression using rank as the independent variable and the rates as dependent variables, we obtain a first-order regression coefficient very close to 0. Thus, we can exclude that website popularity plays a role in how users interact with the Privacy Banner.

B. Geographic Region and `Reject All`

We compare the behaviour of users in different regions of the world. As described in Section III, the CMP implements a Privacy Banner that can take two forms, during *Period A* and *Period B*.

In Figure 4 we provide a breakdown by different geographic regions of the world for the two periods. We group countries by continent but partition Europe in two subsets, considering i) the countries that are part of the European Union (EU) where the General Data Protection Regulation (GDPR) is in force, and ii) all the others. We consider the United Kingdom a GDPR-compliant country because it has a nearly identical regulation. To ensure a fair comparison, we show only the regions for which at least 10 websites had 10 interactions or more in both *Period A* and *Period B*. The red bars show the *Reject-Some* rate during *Period A*; and the blue bars during *Period B*. As described in Section III-D, the values of the bars represent the website-wise macro-average of the rate. Thus, each website has the same weight. The vertical black lines indicate the 90% confidence interval for such average.

Starting from *Period A* (red bars), we do not observe significant differences between regions when all users are shown the Privacy Banner as in Figure 1a, as confidence interval bars overlap, with the exception of GDPR-regulated countries and African countries. In all cases, the rate is primarily due to *Mandatory-Only* interactions, while the percentage of *Custom* interactions is negligible (on the order of 0.1-0.2%).

The blue bars in Figure 4 report the *Reject-Some* rate during *Period B* when users from GDPR countries see the Privacy Banner with the additional `Reject All` button as in Figure 1b. In these countries, grouped on the left of the figure, we observe a sharp increase by a factor of four. The *Reject-Some* rate grows from 3.49% to 20.56%. Non-overlapping error bars show this increase is statistically significant. As expected, we do not observe any significant changes for the other geographic regions as users still interact with the first version of the banner. Overall, this figure shows how the design of the Privacy Banner influences users' decisions. When it is as easy to reject cookies as it is to accept them, more than one in five users chooses to reject them. As a consequence of CNIL fines on Google and Facebook, many European websites and CMPs are implementing similar `Reject All` buttons. In general, we can relate these results to the debate about dark patterns [12, 13]. Our measurements confirm how the options present in the Privacy Banner can influence users' choices on cookies and reveal a nearly 5$\times$ increase in users rejecting cookies when only a single click is required. We stress the importance of being able to quantitatively evaluate said figures. It is interesting to observe that the large fraction of users who opt out of cookies with such a Privacy Banner can somehow impact the business of those portals that rely heavily on tracking and behavioural advertising.

Figure 5 further breaks down the above results by showing the *Reject-Some* rate for different countries. To provide a solid picture, we again limit the analysis to the countries for which we record at least 10 websites with at least 10 interactions in both periods – showing the first 15 countries by descending *Reject-Some* rate in *Period B*. The figure confirms the previous

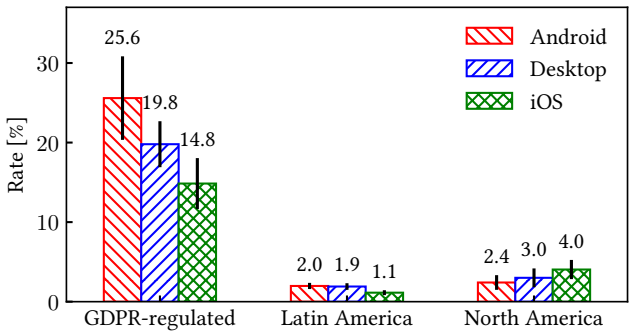


Fig. 6: *Reject-Some* rates according to users' device.

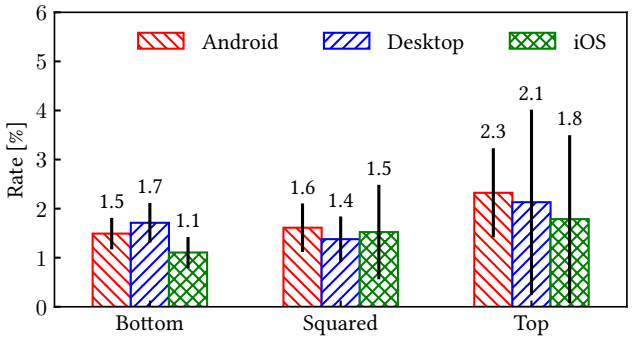


Fig. 7: *Reject-Some* rates according to the position of the screen where the banner appears.

TABLE III: Relative distribution of used devices per region of connection.

Region	Android	iOS	Desktop
GDPR-regulated	41.6%	40.5%	17.9%
Latin America	40.6%	38.9%	20.5%
North America	30.0%	44.8%	25.2%

results. In *Period A*, we do not observe significant differences in the *Reject-Some* rate between GDPR (France, Romania, Germany, Spain and the UK) and non-GDPR countries. Indeed, confidence intervals overlap. In *Period B*, conversely, with the insertion of the additional `Reject All` button, we observe a significant increase in all GDPR-regulated countries, from a $\sim 3.5\times$ in the UK and Germany ($\sim 6\%$ to $\sim 23\%$) to more than $7\times$ in Portugal (2.14% to 15.34%). As expected, there are no significant variations in other, non GDPR-regulated countries.

C. User device type

We now move on to analyze the differences between users browsing the Web with different types of devices. To this end, we categorize each interaction based on the client-side `User-Agent` HTTP header, to obtain the operating system (OS) of the user's device. Considering that the experience of navigating websites is not greatly affected by OS when using a PC, we group Windows, Mac OS, Linux and other operating systems under the same *Desktop* category. Conversely, we divide the mobile landscape into two main major categories: *Android* and *iOS*. Overall, *Desktop*, *iOS*, *Android* represent the 21%, 39% and 40% of the entries, respectively. Other mobile OSes are present in the dataset, but their volume is so low that we neglect them. The region-wise device shares are overall homogeneous across the regions and are reported in Table III.

In Figure 6 we show the *Reject-Some* rate separately by OS. We target *Period B* because our dataset contains the `User-Agent` field only after August 25, 2022. For the non-GDPR regions, we choose North and Latin America as they are the origin of the largest amount of interactions (see Table II). We include a website in the macro-average only if it collected at least 10 interaction in the target (website, region) couple. Overall, the figure confirms the large difference between regions that we discussed in Section IV-B. Moreover, it surprisingly shows that, in countries subject to the GDPR, Android users are more likely to reject some cookies than iOS users – see the first box group in Figure 6. One could argue that iOS users feel safer than Android users due to Apple's efforts to enforce and communicate privacy-preserving technologies on its devices, but the data at our disposal do not allow us to prove any hypothesis about it. This is nonetheless an interesting finding, which offers a stimulating question for future work. The same consideration holds for Latin America, while in North America, the averages have reversed roles. We limit ourselves to observing these figures, while the search for the causes of this behavior requires other data and possibly controlled experiments.

D. Banner size and position

We then investigate whether the shape and dimension of the banner presented to the users impacts their behavior. Figure 7 shows the *Reject-Some* rate according to the type of banner presented to the user. Let us recall that three types of banners are offered by the CMP: a top-screen long and narrow banner (identified as Top), a bottom-screen long and narrow banner (Bottom), and a bottom-left square banner (Squared). Here we limit the analysis to *Period B* and areas that are not regulated by the GDPR and for which we have a larger number of interactions. In fact, the websites that use the Squared banner account for about 5%. Thus, they received a number of interactions by GDPR countries that does not allow to draw solid conclusions.

Overall, Figure 7 seems to suggest there are not significant differences in the way users interact with the banner with respect to its shape and position. Few websites implement the Top banner, resulting in a large confidence interval. Unfortunately, our data do not allow to track the behaviour and the volume of users that *neglect* the Privacy Banner – i.e., do not interact at all with it. Thus, we cannot measure whether the fraction of users interacting over the total visitors differs according to the different position or shape of the banner.

TABLE IV: Breakdown of partial accept among categories of cookies.

Category of Cookies	Acceptance Rate
Necessary	$100.00 \pm 0.00\%$
Statistics	$58.42 \pm 3.19\%$
Preferences	$47.14 \pm 3.23\%$
Marketing	$22.10 \pm 2.68\%$

TABLE V: Number interactions related to users clicking or not on the Cookie Policy (CP) and the Privacy Policy (PP). The last column indicates, the *Reject-Some* rate for the given set of interactions.

PP clicks	CP clicks	Interactions	*Reject-Some* rate
Yes	Yes	349	$7.45 \pm 2.75\%$
Yes	No	944	$6.04 \pm 1.52\%$
No	Yes	1 176	$3.57 \pm 1.06\%$
No	No	1 011 737	$1.01 \pm 0.02\%$

E. Other behaviours

As observed in Section IV-B, users accessing the `Custom Permissions` screen are in the order of a few percentage points. This confirms that the majority of users do not bother taking precautions for their privacy if this requires more than one click. In this section, we characterize the behavior of users dealing with advanced options.

1) Cherry-picking cookies: By clicking on the `Custom Permissions` button, users are offered the possibility to give separate consent for different types of cookies. Out of 4 M user interactions of our dataset, only 647 times users customized consent for different cookie categories – i.e., provided a *Custom* consent. For completeness, in Table IV, we show the acceptance rate for each category, uniquely for the 647 entries that correspond to *Custom* consent. To evaluate confidence interval, we consider that the proportion of users that accept a category of Cookies (e.g., Statistics) is an unbiased estimator of the probability p of a Bernoulli random variable. Assuming that all the interactions are independent repetitions of such random variable, we obtain the number of successes of a binomial random variable. We thus use binomial proportion confidence intervals, with a confidence level of 90%.

Necessary cookies, represented in the first row of the table, are mandatory and, therefore, cannot be disabled by the user as they are required for website operation. Statistics cookies are the most accepted (58% of cases). These cookies are related to analytics services that account for the number of accesses to the website and monitor performance. Preference cookies, used to recognize users when they return to the website, are accepted to a similar extent (47%). Finally, Marketing cookies are most often rejected. Only 22% of users accepted them. These cookies include web trackers and advertising platforms. Users tend to avoid them, and we can guess that they are perceived as the most privacy intrusive.

2) Visualizing policies: We finally quantify the number of users who access the text of the policies regulating the use of personal data in a website. Indeed, websites must offer the possibility to access this information, and the CMP includes links to Cookie and Privacy Policies. Unless the website implements some customization, the Cookie Policy includes a brief explanation on the concept of cookie, information on the categories of cookies collected by the CMP (Necessary, Statistics, Preferences, Marketing) and their purpose. The Cookie Policy is presented as a small pop up (305 word in its default formulation, in English) and the users do not leave the page they are visiting. Conversely, clicking on the Privacy Policy opens a new webpage which can be either hosted on the website or served by the CMP. The Privacy Policy contains information about the use of personal data at large, of which the cookies represent only a subsection. The policy includes, among the rest, information about the purpose of data collection, the parties with which said data might be shared, the retention policy of the data, etc.

Our dataset records all clicks on the Cookie Policy. Those on the Privacy Policy are tracked only if the policy is hosted by the CMP, so we restrict our analysis to approximately one-fourth of the total interactions. Again, due to the low number of interactions of this type, we do not show the website-wise macro-average but provide the overall numbers directly in Table V, while the confidence interval are again calculated using binomial proportion. The number of interactions in which a user either clicks on at least one of the links is very low: 2 469, 0.24% of the total. Users who decide to read (or at least visualize) the policies appear more careful about their privacy: those who click on both policies record a 7.45% *Reject-Some* rate, while users who do not visualize any account for a value of only 1.01%. Although we cannot prove that the *Reject-Some* rate increases *because* users read the policies, there is at least a sizeable correlation between users' interest in the policies and their unconditional *Accepted-All* rate.

V. Limitations and Conclusions

A. Limitations

As the previous sections pointed out, in this paper, we had the unique possibility to analyze data from a medium-sized CMP. This offered us the opportunity to work with a large amount of data and provide solid results, but it came with some limitations.

First, the information related to the user is limited by design. We did not design a data collection to observe users' behaviour specifically, but we collected data of generic users on generic websites on the Internet. We opted to record interactions with the Privacy Banner without asking for additional information. This allows us to easily scale the measurements to all users visiting the CMP's customer websites but limits the information at our disposal. In particular, we cannot stratify by users' characteristics (age, gender, or educational level). To obtain this kind of data, much costlier controlled experiments should

be set up, with users explicitly approving the collection of such sensitive data. We only infer the user geo-location from the client IP /24 subnet (which we sanitized to protect the user's anonymity) or from the user agent string in the HTTP requests.

The most notable limitation is that we can only collect information on users that interacted with the privacy banner offered by the CMP. By design, we cannot collect any information if users did not interact with the banner, in accordance with regulations in force (GDPR among all) that forbid such recording. This prevents us from studying the fraction of users *not* interacting with the banner, which we suppose is not negligible. Again, a controlled experiment should be implemented to obtain such data.

At last, by design, the dataset does not provide a unique identifier for every user interacting with the banner. This is intended to protect users' privacy but limits us in observing the consistency of users' choices across websites. Thus, we cannot provide a user-centric analysis, but we are limited to an interaction-centric analysis.

B. Conclusions

In this work, we analyzed how users interact with the Privacy Banner and how different factors impact their behaviour. Thanks to a dataset containing millions of interactions with the Privacy Banners present on hundreds of websites, we observed which factors impacted the users' actions. These include characteristics of the users (their country, device, visited website) and of the banner itself (position, options offered). We showed that, when offered balanced options to accept or reject the cookie usage, the fraction of users that rejected grows by a factor of $\sim 5\times$ than when the rejection requires more than one click. Regulators have started considering this aspect when proposing solutions to enhance users' privacy on the Web. Among other factors that impact the users' choice, we found that Android users exhibit a significantly higher rejection rate than iOS users. This is particularly evident in GDPR-regulated countries. This may stimulate further investigations, likely in controlled environments.

In general, we believe this paper can foster the discussion on the long-time impact of regulation and public opinion in the field of privacy on the Web, and we hope to open directions for stimulating future work in this field.

References

[1] N. Nikiforakis, A. Kapravelos, W. Joosen, C. Kruegel, F. Piessens, and G. Vigna, "Cookieless monster: Exploring the ecosystem of web-based device fingerprinting," in *2013 IEEE Symposium on Security and Privacy*, pp. 541–555, IEEE, 2013.

[2] E. Papadogiannakis, P. Papadopoulos, N. Kourtellis, and E. P. Markatos, "User tracking in the post-cookie era: How websites bypass gdpr consent to track users," in *Proceedings of the Web Conference 2021*, WWW '21, (New York, NY, USA), p. 2130–2141, Association for Computing Machinery, 2021.

[3] A. Cahn, S. Alfeld, P. Barford, and S. Muthukrishnan, "An empirical study of web cookies," in *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, (Republic and Canton of Geneva, CHE), p. 891–901, International World Wide Web Conferences Steering Committee, 2016.

[4] J. R. Mayer and J. C. Mitchell, "Third-party web tracking: Policy and technology," in *2012 IEEE Symposium on Security and Privacy*, pp. 413–427, 2012.

[5] F. Roesner, T. Kohno, and D. Wetherall, "Detecting and defending against Third-Party tracking on the web," in *9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12)*, (San Jose, CA), pp. 155–168, USENIX Association, Apr. 2012.

[6] S. Englehardt, D. Reisman, C. Eubank, P. Zimmerman, J. Mayer, A. Narayanan, and E. W. Felten, "Cookies that give you away: The surveillance implications of web tracking," in *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, (Republic and Canton of Geneva, CHE), p. 289–299, International World Wide Web Conferences Steering Committee, 2015.

[7] N. Samarasinghe, A. Adhikari, M. Mannan, and A. Youssef, "Et tu, brute? privacy analysis of government websites and mobile apps," in *Proceedings of the ACM Web Conference 2022*, WWW '22, (New York, NY, USA), p. 564–575, Association for Computing Machinery, 2022.

[8] T. Bujlow, V. Carela-Español, J. Sole-Pareta, and P. Barlet-Ros, "A survey on web tracking: Mechanisms, implications, and defenses," *Proceedings of the IEEE*, vol. 105, no. 8, pp. 1476–1510, 2017.

[9] Council of European Union, "Directive 2009/136/EC amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws." http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32009L0136 (Last accessed September 6, 2021), 2009.

[10] European Parliament and Council of European Union, "Directive 95/46/EC. General Data Protection Regulation." http://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf (Last accessed September 6, 2021), 2016.

[11] M. Hils, D. W. Woods, and R. Böhme, "Measuring the emergence of consent management on the web," in *Proceedings of the ACM Internet Measurement Conference*, pp. 317–332, 2020.

[12] C. Matte, N. Bielova, and C. Santos, "Do cookie banners respect my choice? : Measuring legal compliance of

banners from iab europe's transparency and consent framework," in *2020 IEEE Symposium on Security and Privacy (SP)*, pp. 791–809, 2020.

[13] T. H. Soe, O. E. Nordberg, F. Guribye, and M. Slavkovik, "Circumvention by design - dark patterns in cookie consent for online news outlets," in *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, NordiCHI '20, (New York, NY, USA), Association for Computing Machinery, 2020.

[14] R. v. Eijk, H. Asghari, P. Winter, and A. Narayanan, "The impact of user location on cookie notices (inside and outside of the european union)," in *Workshop on Technology and Consumer Protection (ConPro'19)*, 2019.

[15] M. Trevisan, S. Traverso, E. Bassi, and M. Mellia, "4 years of eu cookie law: Results and lessons learned.," *Proc. Priv. Enhancing Technol.*, vol. 2019, no. 2, pp. 126–145, 2019.

[16] I. Sanchez-Rola, M. Dell'Amico, P. Kotzias, D. Balzarotti, L. Bilge, P.-A. Vervier, and I. Santos, "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, (New York, NY, USA), p. 340–351, Association for Computing Machinery, 2019.

[17] A. Dabrowski, G. Merzdovnik, J. Ullrich, G. Sendera, and E. Weippl, "Measuring Cookies and Web Privacy in a Post-GDPR World," in *Passive and Active Measurement* (D. Choffnes and M. Barcellos, eds.), (Cham), pp. 258–270, Springer International Publishing, 2019.

[18] T. Linden, R. Khandelwal, H. Harkous, and K. Fawaz, "The privacy policy landscape after the gdpr," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 1, pp. 47–64, 2020.

[19] M. Degeling, C. Utz, C. Lentzsch, H. Hosseini, F. Schaub, and T. Holz, "We value your privacy... now take some cookies," *Informatik Spektrum*, vol. 42, no. 5, pp. 345–346, 2019.

[20] M. Kretschmer, J. Pennekamp, and K. Wehrle, "Cookie banners and privacy policies: Measuring the impact of the gdpr on the web," *ACM Trans. Web*, vol. 15, jul 2021.

[21] H. Habib, S. Pearman, J. Wang, Y. Zou, A. Acquisti, L. F. Cranor, N. Sadeh, and F. Schaub, ""it's a scavenger hunt": Usability of websites' opt-out and data deletion choices," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, (New York, NY, USA), p. 1–12, Association for Computing Machinery, 2020.

[22] N. Jha, M. Trevisan, L. Vassio, and M. Mellia, "The internet with privacy policies: Measuring the web upon consent," *ACM Transactions on the Web (TWEB)*, vol. 16, no. 3, pp. 1–24, 2022.

[23] Deloitte, "Cookie Benchmark Study." https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/risk/deloitte-nl-risk-cookie-benchmark-study.pdf (Last accessed September 6, 2021), 2020.

[24] J. M. Bauer, R. Bergstrøm, and R. Foss-Madsen, "Are you sure, you want a cookie?–the effects of choice architecture on users' decisions about sharing private online data," *Computers in Human Behavior*, vol. 120, p. 106729, 2021.

[25] P. Hausner and M. Gertz, "Dark patterns in the interaction with cookie banners," *arXiv preprint arXiv:2103.14956*, 2021.

[26] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "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*, CHI '20, (New York, NY, USA), p. 1–13, Association for Computing Machinery, 2020.

[27] V. B. R and R. P. N, "Testing the effect of the cookie banners on behaviour," no. LF-NA-28287-EN-N, 2016.

[28] C. Utz, M. Degeling, S. Fahl, F. Schaub, and T. Holz, "(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, (New York, NY, USA), p. 973–990, Association for Computing Machinery, 2019.

[29] O. Kulyk, W. Rafnsson, I. M. Borberg, and R. H. Pedersen, ""so i sold my soul": Effects of dark patterns in cookie notices on end-user behavior and perceptions," in *Proceedings of 2022 Symposium on Usable Security and Privacy*, Internet society, 2022.

[30] A. K. Singh, N. Upadhyaya, A. Seth, X. Hu, N. Sastry, and M. Mondal, "What cookie consent notices do users prefer: A study in the wild," in *Proceedings of the 2022 European Symposium on Usable Security*, pp. 28–39, 2022.

[31] T. Vila, R. Greenstadt, and D. Molnar, "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*, pp. 403–407, 2003.

[32] J. Grossklags and N. Good, "Empirical studies on software notices to inform policy makers and usability designers," in *International Conference on Financial Cryptography and Data Security*, pp. 341–355, Springer, 2007.

[33] L. M. Coventry, D. Jeske, J. M. Blythe, J. Turland, and P. Briggs, "Personality and social framing in privacy decision-making: A study on cookie acceptance," *Frontiers in psychology*, vol. 7, p. 1341, 2016.

[34] S. Barth and M. D. de Jong, "The privacy paradox – investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review," *Telematics and Informatics*, vol. 34, no. 7, pp. 1038–1058, 2017.