Intractable Cookie Crumbs: Unveiling the Nexus of Stateful Banner Interaction and Tracking Cookies

Ali Rasaii
Max Planck Institute for Informatics

Ha Dao
Max Planck Institute for Informatics

Anja Feldmann
Max Planck Institute for Informatics

Mohammadmahdi Javid
Max Planck Institute for Informatics

Oliver Gasser
IPinfo

Devashish Gosain
IIT Bombay

Abstract

In response to the ePrivacy Directive and the consent requirements introduced by the GDPR, websites began deploying consent banners to obtain user permission for data collection and processing. However, due to shared third-party services and technical loopholes, non-consensual cross-site tracking can still occur. In fact, contrary to user expectations of seemingly isolated consent, a user's decision on one website may affect tracking behavior on others.

In this study, we investigate the technical and behavioral mechanisms behind these discrepancies. Specifically, we disclose a persistent tracking mechanism exploiting web cookies. These cookies, which we refer to as *intractable*, are initially set on websites with accepted banners, persist in the browser, and are subsequently sent to trackers before the user provides explicit consent on other websites. To meticulously analyze this covert tracking behavior, we conduct an extensive measurement study performing stateful crawls on over 20k domains from the Tranco top list, strategically accepting banners in the first half of domains and measuring intractable cookies in the second half. Our findings reveal that around 50% of websites send at least one intractable cookie, with the majority set to expire after more than 10 days. In addition, enabling the Global Privacy Control (GPC) signal initially reduces the number of intractable cookies by 30% on average, with a further 32% reduction possible on subsequent visits by rejecting the banners. Moreover, websites with Consent Management Platform (CMP) banners, on average, send 6.9 times more intractable cookies compared to those with native banners. Our research further reveals that even if users reject all other banners, they still receive a large number of intractable cookies set by websites with cookie paywalls. Additionally, our measurement on the partitioned cookies—cookies that are restricted to the top-level site and thus mitigate cross-site tracking—shows that only 1.3% of tracking cookies are marked as such, indicating their minimal impact on cross-site tracking via intractable cookies.

Keywords

Privacy Regulation, GDPR, Tracking Cookies, Web Tracking, Cookie Banner, Intractable Cookies, Cross-Site Tracking

Proceedings on Privacy Enhancing Technologies YYYY(X), 1–17

https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

In the digital age, online tracking mechanisms have become a common aspect of Internet use [3, 7, 30, 40, 57]. These mechanisms—such as third-party tracking cookies—form the foundation of data-driven marketing and advertising strategies. Statistics show that major tech companies depend heavily on online advertising [6, 21], employing various tracking techniques to monitor user activity and optimize personalized ads. The opaque nature of data-handling practices has led to a growing demand for transparency and control over personal data among privacy advocates [5, 50, 69]. In response, legislative bodies have introduced strict data protection laws, like the ePrivacy Directive [28] and the General Data Protection Regulation (GDPR) [15] in the European Union and the California Consumer Privacy Act (CCPA) [13] in California. These laws aim to limit unrestricted user data collection and enforce a more transparent consent process.

Consequently, a significant shift toward preserving user privacy can be observed. Research indicates a reduction in third-party tracking and increased visibility of privacy policies within the EU [18, 19, 48, 72]. To comply with the ePrivacy Directive's mandate requiring prior consent for storing or accessing non-essential information on a user's device, websites increasingly deploy cookie banners to inform users and obtain their consent [19, 65]. The GDPR further strengthened and enforced the standards and legal criteria for valid consent established under the ePrivacy Directive. Despite these advancements, previous studies [52, 65, 73] demonstrate that websites continue to set tracking cookies extensively, even without explicit user consent. While the mere act of placing such cookies is not necessarily a violation of privacy regulations—and, in some cases, may be justified under Article 6 GDPR—the fact that most of these cookies originate from well-known tracking companies raises concerns about the alignment of such practices with the principles of transparency, purpose limitation, and user autonomy.

Furthermore, the design and implementation of consent banners have been criticized for employing manipulative tactics that skew user behavior, often leading to uninformed or coerced consent decisions [47, 51, 71]. For instance, users may accept cookie banners when there is no easy refusal option, resulting in the setting of tracking cookies. These unintentionally accepted cookies may increase the likelihood of being tracked during future web browsing. The stateful cross-site reuse of tracking cookies before obtaining user consent is a previously under-explored and unmeasured threat. Exploring cookie banners and their impact in a stateful manner is

therefore crucial, as both browsing and tracking unfold over time, often involving interconnections across multiple sites and entities.

In this paper, we uncover a persistent inter-domain tracking mechanism via web cookies. Specifically, we find that numerous tracking cookies, *initially set on a website with an accepted banner, continue to be transmitted to tracker domains even before users interact with the rejectable banners* on other websites. Unlike previous studies that focused on the *stateless deployment* of cookies on websites, we conduct a *stateful* banner interaction across websites, demonstrating how seemingly opposing decisions on one website (with accepted banner) can influence tracking behavior on subsequent websites (with rejected banner), resulting in the *transmission* of cookies to trackers. This not only violates privacy regulations such as the GDPR by undermining the overall effectiveness of banners as the primary consent mechanism, but also creates a false sense of privacy when users reject the banners. Throughout this paper, we refer to these cookies as **intractable cookies** (see § 3).

To substantiate our findings, we conduct two measurement campaigns using BannerClick [65], a tool designed to automatically detect and interact with cookie banners. We improved its performance, particularly by increasing its rejection accuracy from 87% to 99% (see § 4). In the first campaign, we crawl the top 20,000 websites from Tranco, accepting cookie banners in the first half and measuring the number of intractable cookies on successfully rejected domains in the second half. For the second campaign, we randomly select and shuffle 20k sites from Tranco's top 50k, following a similar approach. Additionally, we perform both runs in reverse order to gain further insights. Overall, our main findings can be summarized as follows:

- We find that nearly 50% of websites send at least one intractable cookie to third-party tracking domains before obtaining explicit user consent (see § 5.1).
- Regarding the effect of banner interaction, we see no immediate changes on intractable cookies after rejecting banners. However, on average, 25% of intractable cookies are not sent after reloading the webpage with the rejected banner (see § 5.2).
- Furthermore, our measurements show that enabling the Global Privacy Control (GPC) signal in the browser can initially reduce intractable cookies by an average of 30%. An additional 32% reduction is achievable on subsequent visits by also rejecting cookie banners (see § 5.3).
- We observe that more popular Tranco websites send fewer intractable cookies compared to less popular ones. Specifically, the top 50 websites send, on average, zero intractable cookies, while the top 10k websites send 25 intractable cookies (see § 5.4).
- We note that websites using CMP banners send more than 6.9 times as many intractable cookies compared to those using native banners. Moreover, out of 5,915 accepted domains, 90 websites with cookie paywalls are responsible for setting more than 35% of intractable cookies (see § 5.5).
- We also observe ≈ 60% of intractable cookies have an expiration time of at least 10 days, highlighting their persistence. In addition, around 90% of intractable cookies are set (or refreshed) by at most 1% of accepted websites (see § 5.6).
- Our analysis shows that, on average, each domain has 3.42 different trackers, with each tracker receiving an average of 7.3 intractable cookies. We also verified that the top 20 trackers are indeed well-known tracking companies (see § 5.7).
- Additionally, to assess the impact of partitioned cookies—whose transmission is restricted to their *setter website*— in mitigating intractable cookies, we conduct a separate measurement using Chrome. The results show that only 1.3% of all unique tracking cookies are partitioned, with more than half accompanied by non-partitioned cookies from the same tracker domain (see § 5.8).

References	Automated	Reject Coverage	Stateful	Sent Cookie
Englehardt et al. [24]	✓	N/A	✓	✗
Trevisan et al. [73]	✓	CMP	✗	✗
Matte et al. [52]	✗	CMP	✗	✗
Jha et al. [43]	✓	N/A	✗	✗
Smith et al. [70]	✓	CMP	✗	✗
Rasaii et al. [65]	✓	87%	✗	✗
Our work	✓	99%	✓	✓

Table 1: Overview of the closest previous studies on the misbehavior of tracking cookies. Our study is conducted in a fully automated manner, capable of rejecting 99% of all banners. Furthermore, it is the first study to statefully measure the act of sending cookies to the tracker rather than their setting in the browsers.

Finally, our findings reveal a persistent gap between privacy regulations, such as the GDPR and ePrivacy Directive, and their technical implementation. In § 6, we examine how fragmented interpretations and unclear accountability contribute to this gap, and introduce our browser-integrated approach as a potential solution.

2 BACKGROUND AND RELATED WORK

Statistics [6, 21] show that major tech companies rely heavily on online advertising, with Google generating 77% and Facebook 98.4% of their revenue from ads. These companies employ various tracking techniques, e.g., web cookies, to collect users' online activity and deliver targeted ads while optimizing ad recommendations. Subsequently, previous studies [5, 50, 69] show that people are increasingly concerned about how their personal information is being collected and used by companies. In recent years, several data protection laws have been enacted to regulate the use of web cookies and other tracking and profiling techniques, such as the ePrivacy Directive [28] and the General Data Protection Regulation (GDPR) [15] in the European Union and the California Consumer Privacy Act (CCPA) [13] in California.

The GDPR requires websites to obtain user consent before collecting or processing personal data, unless the processing is justified by another legal basis—such as being strictly necessary for the performance of a service requested by the user (Article 6). As a result, websites increasingly rely on cookie banners to inform users about data practices and to obtain consent. Many outsource this functionality to Consent Management Platforms (CMPs)—third-party services that provide ready-to-use, configurable consent interfaces. A widely adopted framework for implementing these banners is the IAB Europe Transparency and Consent Framework (TCF) [39], which standardizes how CMPs operate and transmit consent signals. While the TCF is presented as a GDPR-compliant solution to facilitate user consent and regulatory compliance, it originates from

IAB Europe—an industry group representing the online advertising sector. As such, CMPs built on the TCF may prioritize maximizing consent rates, aligning with the interests of advertisers rather than promoting user privacy.

Numerous studies have examined compliance issues, particularly identifying potential legal violations in cookie banner implementation and consent storage. While these studies show that the GDPR has led to a per-site reduction in third-party tracking and improved the visibility of privacy policies and cookie banners across the European Union, violations persist [18, 19, 48, 72]. In addition, some studies have evaluated the influence of consent banner design on user behavior, specifically their acceptance or denial of consent [8, 47, 51, 71]. Santos et al. [68] analyzed the clarity of cookie banners and found that 61% of them employed vague language, failing to specify privacy practices adequately. Utz et al. [76] explored additional factors influencing user consent, such as banner placement, and reported significant impacts on consent decisions based on these elements. Additionally, Nouwens et al. [59] showed that merely removing the opt-out button from the first layer of banners increases consent rates by approximately 23%. Overall, these studies consistently identified interface interference as a key factor that significantly influences how users interact with banners.

Finally, many studies specifically analyzed the techniques and prevalence of tracking cookie usage in web tracking. Table 1 compares some of the most relevant studies [24, 43, 52, 65, 70, 73] with our work based on key distinguishing factors, such as whether they were conducted manually or through automation, their rejection coverage, whether they employ stateful or stateless measurements, and whether they focus on the setting of cookies on browsers or sending them to tracker domains.

Englehardt et al. [24] conducted one of the first fully automated, large-scale studies on tracking cookies. They introduced *OpenWPM*, a measurement platform capable of collecting HTTP requests and responses, JavaScript calls, and script files. Their analysis included an ID cookie detection method to uncover cookie syncing across sites. However, even though they employed a stateful approach, their analysis relied on identifying user IDs previously set in cookies and later used in referer headers and request URLs; they did not examine the `Cookie` HTTP header to observe cookies being sent directly to servers. Nevertheless, the study was conducted in January 2016, prior to the adoption of the GDPR, and it did not explicitly discuss the implications of its findings in the context of existing regulations at the time, such as the ePrivacy Directive, nor did it consider the effect of banner interaction. Trevisan et al. [73] developed CookieCheck, a tool that visits websites as a new user and analyzes cookies placed in the browser. It focused solely on profiling cookies set by CMPs that violate the ePrivacy Directive before any user consent is given. Matte et al. [52] conducted semi-automatic crawl campaigns to detect suspected GDPR and ePrivacy Directive violations in banners based on the Transparency and Consent Framework developed by IAB Europe. Jha et al. [43] attempted to interact with cookie banners in an automated manner to observe differences in the cookies set. However, their work focused solely on banner acceptance in a stateless manner. Smith et al. [70] specifically investigated the placement of tracking cookies under the guise of legitimate interest by CMPs, as well as their compliance with properly transmitting users' choices through TCF

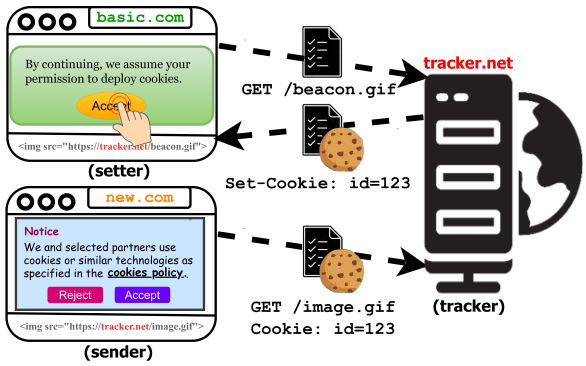


Figure 1: Overview of a scenario illustrating the interconnections among the four main entities involved in intractable cookie transmission: *setter website* with an accepted banner, *tracker domain, sender website* with a rejectable cookie banner, and the browser.

consent strings. Finally, Rasaii et al. [65] developed BannerClick, a tool capable of both accepting and rejecting consent banners. They conducted a comprehensive measurement study on Tranco's Top 10K sites, analyzing cookies deployment in a stateless manner.

In this research, we investigate how cookies set on a website with an accepted banner may contribute to user tracking on subsequent visited websites before any consent is given. Among the available tools capable of interacting with banners [10, 58], we use an improved version of BannerClick due to its higher rejection coverage and integration with OpenWPM. This integration enables us to perform a combination of stateful and stateless crawls, capturing all the data required. We then examine the implications of our findings concerning privacy regulations and potential violations. To the best of our knowledge, this specific form of user tracking via web cookies has not been previously explored in the literature.

3 INTRACTABLE COOKIES

In this section, we detail the nature of intractable cookies and the entities involved, as illustrated in Figure 1.

In the depicted scenario, a user initially visits `basic.com` and "accepts" the cookie banner. This might result in setting new cookies in the browser. For example, the browser starts sending HTTP requests to load a third-party resource, `beacon.gif`, from `tracker.net`, which sets a cookie (*id*=123). We refer to `basic.com` as the *setter website*, as it establishes the context for the initial cookie-setting in the user's browser, and to `tracker.net` as the *tracker domain*, as it tracks the user. Later, the user accesses `new.com`, which shows a banner with the "reject" option.[1] In this scenario, even though the user has not yet interacted with the banner to explicitly consent to the cookies, during the rendering of the webpage, the browser might send an HTTP request to load third-party resources. For example, `new.com` might embed a resource (`image.gif`) from `tracker.net`,

[1]We focus on websites with rejectable banners when measuring intractable cookies on *sender websites*. This is because banners—and the associated tracking behavior—are directly shaped by privacy regulations such as the GDPR. A core requirement of these regulations is that users must be given the ability to reject cookies. Consequently, websites whose banners lack a reject option are excluded from our study.

resulting in sending an HTTP request along with the previously stored cookie (*id*=123) back to the *tracker domain*. In this context, `new.com` is referred to as the *sender website*, as it leads to sending the cookie to the *tracker domain*. Ultimately, we define **intractable cookies** as those that are initially set on a website where the user accepts the banner (i.e., , the *setter website*), and are later sent by a *sender website* to the *tracker domain* before any banner interaction.

It should be noted that existing research primarily assesses the setting of tracking cookies prior to explicit consent in a stateless manner. However, in our study, we introduce intractable cookies to address two key factors overlooked by previous work. First, our measurements are conducted in a stateful manner—an essential perspective, as internet browsing and tracking inherently occur across sessions and time, with websites often interacting with one another. Second, we focus on the transmission of cookies rather than solely their deployment in browsers. While the GDPR governs the processing of personal data and requires a lawful basis such as consent, the act of setting a cookie may not, in itself, violate the GDPR unless personal data is involved. Though under the ePrivacy Directive (Article 5(3)), storing or accessing tracking cookies on a user's device without prior consent constitutes a violation, regardless of whether the data qualifies as personal. In addition, as demonstrated in Appendix C, the setting of intractable cookies by *sender websites*[2] is relatively uncommon. In many cases, intractable cookies are transmitted to third-party trackers without being explicitly set again. Therefore, treating a website as a single, isolated entity and analyzing the deployment of tracking cookies in a stateless manner may fail to capture the full extent of unwanted tracking practices in the wild and subsequently hinder regulators from crafting the most effective regulatory measures.

In § 6, we elaborate more on the relation between the intractable cookies and privacy regulations, as well as possible mitigation approaches like partitioned cookies [31] and their challenges.

4 METHODS

We now describe our data collection approach, our crawling setup and tools, and the cookie classification method, followed by a discussion of measurement limitations.

4.1 Data Collection

To overcome the shortcomings of existing popular site lists—such as instability, unreachable domains, and susceptibility to manipulation by adversaries [49]—we utilize the Tranco website popularity ranking.[3] Our target list consists of the landing pages of top-ranked domains. To conduct large-scale automated crawls and collect data for analysis, we use a modified version of BannerClick [65], a tool that effectively detects and interacts with cookie banners, and specifically identifies CMP and cookie paywall banners. BannerClick is built on top of OpenWPM [24] (version 0.26.0), which uses Firefox v121.0 with TCP disabled [56]. OpenWPM is widely used in privacy measurement studies and enables the collection of cookies, JavaScript function calls, web resources, and HTTP/HTTPS requests and responses for each site.

[2]For instance, visiting `new.com` might also set the cookie (*id*=123), which could be considered more detrimental than merely sending it.

[3]This list, generated on 07 December 2023 with ID K2NZW, is available at: https://tranco-list.eu/list/K2NZW.

Run	#Crawls	#Acc	#Rej	Date	Duration
Popularity	20k	3,034	2,379	Jul 2024	14 days
Popularity-Reverse	20k	3,060	2,424	Jul 2024	14 days
Random	20k	2,933	2,578	Jul 2024	15 days
Random-Reverse	20k	2,983	2,543	Jul 2024	15 days
RandComb	–	5,916	5,121	–	–
GPC-Enabled	5,121	–	4,947	Aug 2024	5 days
Partitioned Cookies	6,518	4,124	–	Jan 2025	11 days

Table 2: Overview of different measurement types. The columns respectively show the total number of unique domains that are crawled, accepted, and rejected, as well as the date and duration of each run. Note that the *RandComb* run is a combination of the *Random* and *Random-Reverse* runs.

Figure 2 depicts the overview of our methodology setup. Each measurement consists of two phases: stateful and stateless. In the stateful phase, BannerClick crawls the first half of the target list. Upon successful *acceptance* of a banner, we aggregate the corresponding cookies, along with their *setter website*, into a database called *Cookie Jar*. At the end of this phase, we store the browser profile to use as the base profile in the next phase.

Next, BannerClick crawls the second half of the target list in a stateless manner. For each crawl, it loads the final browser profile from the first phase, making the crawl stateful with respect to the accepted domains.[4] After accessing the domain, it waits for 30 seconds and collects all cookies sent through HTTP requests. These cookies are referred to as *Sent Cookies* and are later compared with those in the *Cookie Jar* to detect intractable cookies. In this step, BannerClick interacts with the banner in two separate iterations: one for rejecting and one for accepting. Following each interaction, it waits 10 seconds to capture any immediate changes. Then, to assess if rejection reduces the number of intractable cookies on subsequent visits, it reloads the webpage and waits another 30 seconds. The reload event is conducted in a stateful manner, with browser caching disabled to ensure any inconsistencies are related to banner rejection.

We conduct two main measurement campaigns to analyze intractable cookies, along with two additional runs to investigate the effects of the GPC signal and partitioned cookies. All runs are executed from a server located in the EU, and each crawl within a run is performed once without repetition. Table 2 provides an overview of these runs:

(1) **Popularity runs**: In this measurement campaign, we use Tranco's top 20k websites, first attempting to "accept" banners from the top 10k websites and then detecting intractable cookies on the bottom 10k. This process is conducted with swapped domain lists to explore the impact of website rankings. We refer to these runs as *Popularity* and *Popularity-Reverse*.
(2) **Random runs**: To further assess the intractable cookies on a randomized selection of domains, we sample and mix up 20k domains from Tranco's top 50k domains list. This shuffled list is then split and examined in two sets, as in the popularity runs. We call these runs *Random* and *Random-Reverse*. Moreover, as

[4]Throughout this paper, "accepted domains" refer to domains whose banners have been successfully accepted; the same applies to "rejected domains".

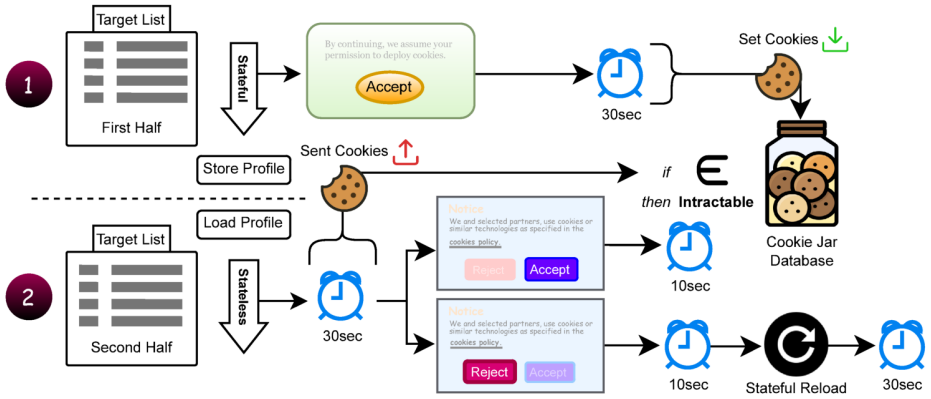


Figure 2: Overview of our methodology: Each run consists of two phases. In the first, banners for half of the target list are accepted in a stateful manner. In the second phase, the browser profile from the first phase is loaded, and two iterations are performed on the remaining domains: one accepting and one rejecting the banners. Cookies set during the first phase on accepted domains and sent to trackers in the second phase before rejecting banners are classified as *intractable cookies*.

shown in § 5.1, both random runs exhibit similar behavior. Thus, for a more comprehensive analysis of the impact of banner interaction, website rank, and banner types on intractable cookies, we use a combined version of them, collectively referred to as the *RandComb* run.

(3) **GPC-Enabled run**: We conduct separate stateless runs to measure the impact of the Global Privacy Control (GPC) signal on preventing intractable cookies. Using browser profiles saved during the stateful phase of the *Random* and *Random-Reverse* runs as the base, BannerClick enables the GPC signal and attempts to reject banners previously identified as rejectable in the random runs (i.e., 5,121 in total). We then combine the results of both runs—similar to the *RandComb* run—and use them for the GPC analysis presented in § 5.3.

(4) **Partitioned Cookies run**: In addition, to assess the prevalence and effectiveness of partitioned cookies in mitigating intractable cookies by limiting their transmission to the *setter websites*, we conduct a separate stateful run using Chrome on the domains with previously detected banners on them in the `RandComb` run (i.e., 6,518 in total), accepting their banners and collecting all cookies set during the visits. Since OpenWPM only supports Firefox and, unlike Chrome, Firefox does not expose the `partitioned_key` field in its cookie storage, we modified `BannerClick` to operate in Chrome's default mode. This setup allows us to collect and measure partitioned cookies independently of OpenWPM. The results of this measurement, along with additional details on the role of partitioned cookies in mitigating cross-site tracking, are presented in § 5.8.

Throughout our analysis, to eliminate the potential impact of differing numbers of accepted domains on the cookies stored in the *Cookie Jar*, we randomly sample and normalize the number of accepted domains when comparing two runs. Specifically, for the run with more accepted domains, we construct a new version of the *Cookie Jar* that includes only cookies whose *setters* belong to a randomly selected subset of accepted domains. The size of this subset matches the number of accepted domains in the run with fewer acceptances. For example, when comparing the *Popularity* and *Popularity-Reverse* runs, we sample 3,034 out of 3,060 accepted domains in the *Popularity-Reverse* run and retain only those cookies in its *Cookie Jar* whose *setter website* is among the sampled domains. This ensures that both *Cookie Jar* instances are derived from the same number of accepted domains, thereby mitigating potential bias when comparing sent cookies.

Crawl Coverage: Out of 20,000 crawls of unique domains in the *Random* run, 16,296 pages were successfully loaded, 692 triggered timeouts, 219 threw exceptions (e.g., due to errors during banner detection or interaction), and 2,783 were completely unreachable. The *Random-Reverse* run exhibits similar behavior.

4.2 Modifications to OpenWPM

We use OpenWPM [24] to conduct a combination of stateful and stateless crawls aiming to observe the interplay of the *setter* and the *sender website* as mentioned earlier. OpenWPM triggers a new event whenever cookies are set, altered, or deleted and stores the corresponding data in the database. We adapted its functionality to ensure that every event involving the addition or updating of a cookie results in overwriting the existing cookie in the browser with a new fixed expiration time (Saturday, 01 Jan 2028, 12:12:12 GMT). This modification is crucial because many cookies might otherwise expire before the completion of our measurements (see § 5.6). Moreover, a cookie can be deleted by overwriting its expiry time before the current time. We identify such cases and allow them to proceed without extending the expiration time. In addition, we frequently store the browser profile to resume from the last saved profile upon unexpected crashes.

Furthermore, we integrated a request parser within OpenWPM that processes HTTP requests by parsing the headers and storing all associated cookies along with the corresponding *sender website*. This enhancement facilitates the comparison of cookies retrieved from HTTP requests on domains where banners are rejected (i.e., *Sent Cookies*) with those stored in *Cookie Jar*.

4.3 Modifications to BannerClick

To cover a broader range of banners in our study, we enhanced BannerClick's accuracy in detecting and interacting with banners, particularly newer types of banners that it was previously unable to handle. In particular, we observed that many banners now provide buttons allowing users to accept (or save) the granular options through the "Settings" layer. Since these granular options typically preselect only essential cookies, clicking these buttons is expected to reject tracking cookies. See Appendix B for an example of such banners. To handle these cases, we modify BannerClick so that if it cannot find the "Reject" button in the banner's main layer, it attempts to click the "Settings" button. If successful, it then attempts to click "Reject All" buttons (or semantically similar ones such as "Decline All," including equivalents in other languages) within the "Settings" layer. If the "Settings" layer remains visible in the viewport, BannerClick then tries to detect and click buttons containing text such as "Confirm," "Save," "Accept Selected," or their semantic or linguistic equivalents.

To evaluate the accuracy of banner rejection after the changes, we randomly selected 1k domains from the top 20k in the Tranco list and attempted to reject their banners. Out of 351 websites with detected banners, BannerClick successfully rejects 263 of them, of which 39 are rejected by clicking on the "Settings" option (20 with preselected options and 19 with a "Reject All" option), while the remaining 224 are directly rejected. We manually reviewed the screenshots and found that only one is a false positive and four are false negatives,[5] resulting in improving the rejection accuracy from 87% [65] to around 99%. Note that among the rejected banners of our sample set, there are no cases where the preselected preferences include non-essential cookies. We publish the modified versions of OpenWPM and BannerClick, as well as analysis scripts and data at `bannerclick.github.io`.

4.4 Cookie Classification

To categorize cookies as first-party or third-party, we utilize the latest Public Suffix List [55] to determine the effective top-level domains (eTLD+1) of both the visited websites and the cookies' `host` attribute. We then compare the domain of each cookie to the domain of its *setter website*. If they match, the cookie is classified as first-party; if not, it is deemed third-party.

Furthermore, we utilize the *justdomains* blocklist [44] to identify *tracker domains*. This list aggregates domain entries from several widely used filter lists, including EasyList, EasyPrivacy, AdGuard, and NoCoin, and has been adopted in prior studies [35, 64, 65]. To obtain the most recent version (i.e., February 2025), we convert the blocklists into the equivalent justdomain lists using the JustDomain converter script [45]. Next, we compare the domains in the lists with the host attribute of the cookies. If there is an exact match or if the host ends with a domain from the list, preceded by a period ('.'), we classify the cookie as a tracking cookie.[6]

Subsequently, we categorize a tracking cookie as *intractable* if it is set on a website with an accepted banner during the first phase of the run (i.e., the stateful phase), and later sent to the tracker from a website in the second phase (i.e., the stateless phase) before its banner is successfully rejected.

4.5 Measurement Limitations

Despite our best efforts to eliminate bias from our measurements, we acknowledge that our study does not fully capture the variety of real-world scenarios users encounter while browsing. First, the individual browsing behaviors are far more complex than the direct crawling of a list of domains. Second, website responses may vary based on several factors, including user activities such as scrolling, and variations in browser settings and capabilities. For example, a study shows that further user interaction and deeper crawling can lead to a 36% rise in the use of tracking technologies like cookies [75]. Prior work also shows that websites can exhibit varying behavior across crawls, even when performed within a short time frame. Moreover, factors such as the user agent and whether the crawler lands on the homepage or an inner page can affect the results [20, 65]. Additionally, our artificial extension of cookie expiration times may introduce some bias; however, as discussed in § 5.6, this effect is minimal.

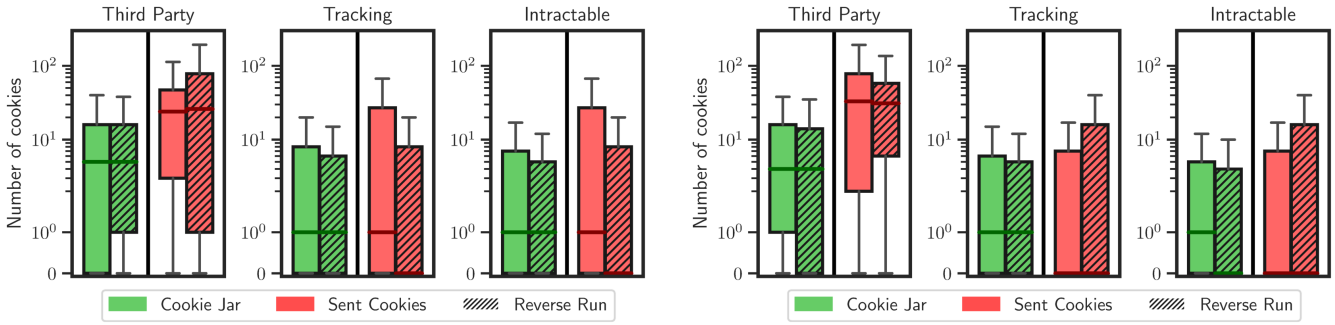
Furthermore, classifying cookies as trackers presents inherent challenges due to the lack of definitive references on their actual usage. Among the many available approaches, each with its own limitations and advantages [14, 34, 57], we identify tracking domains using the justdomains block filter list due to its widespread use and comparability with related studies. However, these lists are crowdsourced, meaning they are continuously updated and maintained by volunteers. As a result, they may overlook certain cases or contain misconfigurations in their exception procedures. For instance, during manual inspection, we identified several third-party cookies classified as intractable cookies, owned by well-known trackers such as `doubleclick.net` and `demdex.net`, that were not included in the final justdomains lists. In particular, we found that `doubleclick.net` used a cookie named `IDE` as an intractable cookie on 2,300 rejected websites. According to Google documentation [32], this cookie is used to record users' interactions with websites' front-end to enable personalized advertising. Moreover, we use a domain-only filter list instead of a full rule-based filter list because our focus is on identifying data transmissions to known *tracker domains*, rather than directly measuring or attributing tracking cookies. While this simplified approach may miss some tracking cookies, it still allows us to capture the broader network interactions with tracker infrastructure.

5 RESULTS

In this section, we present our results and analyze their implications. First, we measure the prevalence of intractable cookies and examine how they are affected by banner interactions and the GPC signal. We then assess how factors such as website ranking and banner type influence the deployment and transmission of intractable cookies. Finally, we investigate their persistence, association with major trackers, and the potential effectiveness of partitioned cookies in mitigating third-party cross-site tracking via intractable cookies.

[5] The false positive occurs when clicking a "Não, ajustar" button that opens the "Settings" layer (i.e., `kinghost.com.br`). False negatives are mainly due to language limitations and uncommon cookie banner designs, such as slide-out panels triggered by the "Settings" button (e.g., `edg.io`).

[6] We iterate over all entries in the filter lists and compare them with the host of cookies using the following condition:
`if host == entry or host.endswith('.' + entry) then True`.



(a) Popularity and Popularity-Reverse runs.

(b) Random and Random-Reverse runs.

Figure 3: Cookie distribution for popularity and random runs.

5.1 Cookie Distribution

Figure 3 shows the overall distribution of third-party, tracking, and intractable cookies across popularity and random runs. Cookies set in accepted domains in the stateful phase (i.e., *Cookie Jar*) and those sent via HTTP headers before the successful rejection of banners in the stateless phase (i.e., *Sent Cookies*) are depicted using green and red boxplots, respectively. Note that, for measuring intractable cookies in *Cookie Jar* (i.e., intractable category green boxplots), we check if a cookie with the same `name` and `host` (the *setter website* might be different) is later sent as intractable cookies. Therefore, when we refer to a cookie in *Cookie Jar* as intractable, it does not imply that it is inherently an intractable cookie. In other words, intractable cookies are essentially third-party tracking cookies that are set with user permission on one website but later sent and propagated covertly without user permission to other websites across stateful browsing sessions, ultimately subverting the core purpose of cookie banners.

In Figure 3a, for tracking cookies, we see nearly identical *Cookie Jar* boxplots (with median one) across both *Popularity* (green) and *Popularity-Reverse* (shaded-green). However, more intractable cookies are sent in the *Popularity* run compared to *Popularity-Reverse* run on average (i.e., the median in the red boxplot is one while it is zero for the shaded-red boxplot). This may indicate that either more popular websites set or the less popular ones send more intractable cookies. We further investigate the impact of website rankings in § 5.4. For the random runs (Figure 3b), the medians of intractable cookies are zero for both, indicating relatively consistent behavior.

In total, it is evident that intractable cookies are common across websites. In all runs, the majority of tracking cookies set in the *Cookie Jar* are identified as intractable, e.g., out of 3,583 unique tracking cookies in the *Popularity* run, 2,131 cookies are later classified as intractable (see Appendix C for details). Note that in our measurements, as detailed in § 4.4, intractable cookies are by definition a subset of detected tracking cookies. Although we consider both first- and third-party cookies when identifying tracking cookies, ≈ 95% of them are third-party cookies in all runs.

To further analyze the intractable cookie distribution, we plot the Empirical Cumulative Distribution Function (ECDF) of intractable cookies for popularity and random runs (see Figure 4). The graphs

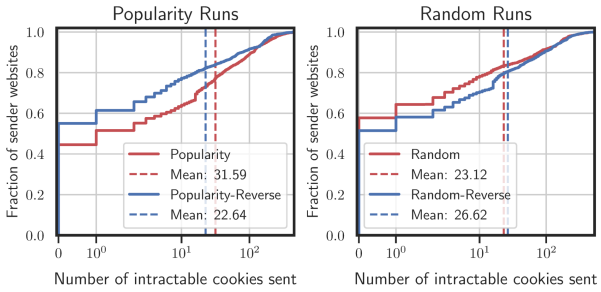


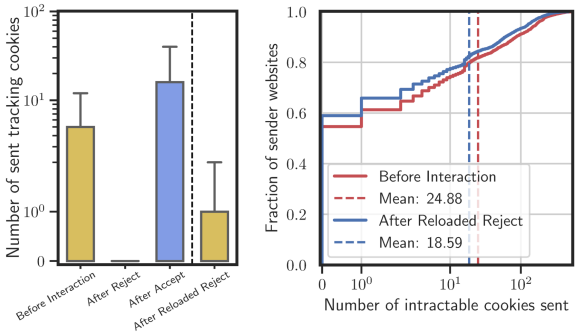
Figure 4: Intractable cookie distribution over websites.

indicate that around 45% to 55% of websites send at least one intractable cookie across all runs. Moreover, we observe on average ≈ 40% more intractable cookies in the *Popularity* run compared to the *Popularity-Reverse* run, while this difference is relatively smaller for random runs, again indicating the possible impact of website ranking on intractable cookies (see § 5.4).

5.2 Impact of Banner Interaction

By definition, intractable cookies are sent to tracker domains prior to any interaction with the banner. We now investigate the possible effect of banner interaction on the later transmission of intractable cookies. As mentioned in § 4, for each domain in the stateless phase, BannerClick performs two separate iterations: one for rejecting the banner and then reloading the webpage, and one for accepting it.

The first three boxplots in Figure 5a show the number of tracking cookies sent upon visiting a webpage that were previously stored in the *Cookie Jar*, corresponding to three stages: before interaction (i.e., intractable cookies), after rejecting, and after accepting the cookie banner for *RandComb* run. The flat line for the "After Reject" box indicates that rejecting the banner does not trigger sending additional tracking cookies previously set. We also find that just a few cookies are explicitly deleted after rejecting the banners. In contrast, accepting the banner triggers the transmission of a new set of cookies (i.e., the blue box), which can be considered a valid action since it reflects user consent to the banner. Additionally, we observe that cookies set after accepting the banner have a median



(a) Impact of Banner Interaction (b) Impact of Reload

Figure 5: Impact of banner interaction and reloading on the number of tracking cookies in Cookie Jar sent. Yellow boxes in (a) are intractable cookies whose ECDF plot is shown in (b) for Before Interaction and After Reloaded Reject.

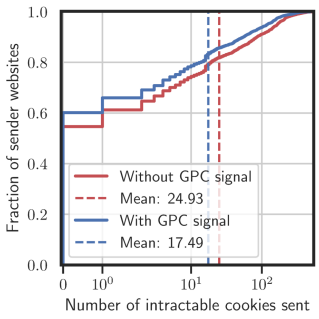


Figure 6: Impact of enabling GPC signal on the number of intractable cookies.

expiration time of 6 months—twice as long as intractable cookies set prior to banner interaction, which have a median of 3 months.

Additionally, the right-most box in Figure 5a shows the number of intractable cookies continuing to be sent after revisiting the rejected website (i.e., a subset of intractable cookies). Compared with the "Before Interaction" box, we see the third quartile dropping from 5 to 1. To further assess the impact of this reload, we plot the ECDF graph of intractable cookies in Figure 5b, comparing the stages before interaction and after reloading the rejected webpage. We observe that, on average, websites stop sending ≈ 25% of intractable cookies after reloading the rejected webpages.

Overall, considering the "After Accept" stage, we observe that initiating tracking after user interaction with banners is technically feasible. However, as the prevalence of intractable cookies indicates, prioritizing tracking before obtaining consent (i.e., the "Before Interaction" stage) remains a common practice on the web.

A Case Study: Through our manual inspection, we observed that interacting with a banner can modify a website's front-end configuration. For instance, rejecting a banner may exclude third-party tracking resources from the HTML source. Conversely, accepting a banner may prompt websites to inject additional third-party resources into the HTML, leading to the transmission of new cookies.

Notably, we found that one of the major entities that govern the website's behavior regarding loading third-party resources is the Consent Management Platform (CMP) (see § 5.5 for more details). For example, `ritzcarlton.com` uses `OneTrust`, one of the most popular CMP [36, 65], to manage user preferences and interactions with the banner. Upon visiting `ritzcarlton.com` for the first time, it renders an iFrame from `demdex.net`, a domain associated with Adobe Audience Manager.[7] In this case, the script inside the iFrame creates new cookies and sends them via XMLHttpRequest API calls to different tracker domains. By default, the browser also sends all previously set cookies along with these requests. Tracker domains can potentially link these cookies together and create a user profile using cookie synchronization techniques [61]. After the user rejects the banner, the CMP stores the user's preference in a cookie, stopping the iFrame from loading and preventing further requests to trackers on subsequent visits to the webpage.

In total, out of the 127,645 intractable cookies detected in the *RandComb* run, around 73% are sent as a result of HTTP requests made by the browser to fetch third-party resources (e.g., img, script, beacon, etc.). The remaining 27% cookies are directly sent via XMLHttpRequest API calls (e.g., *fetch()* method) from script codes.

5.3 Impact of Global Privacy Control

In addition to interacting with cookie banners, other official, standardized mechanisms have also been developed. One of the most notable is Global Privacy Control (GPC) [33], which has recently gained more attention and is now supported by many browsers and extensions.[8] GPC is a browser setting that signals a user's preference not to be tracked. When enabled, the browser informs websites that users do not want their data to be sold or shared. In 2021, the California Attorney General confirmed that businesses must honor the GPC signal as a valid request to opt out of data sales under CCPA [12]. While GDPR does not explicitly mandate GPC, the signal can be interpreted as a withdrawal of consent for tracking, which websites operating under GDPR should honor.

To measure its impact on the intractable cookies, we perform another stateless run as detailed in § 4.1 with the GPC signal enabled. Figure 6 shows the difference in intractable cookies between the *RandComb* run with and without GPC enabled. We observe that enabling GPC reduces intractable cookies by approximately 30% on average, and about 68% of the remaining cookies overlap with those observed in the "After Reloaded Reject" case shown in Figure 5. This indicates that enabling the GPC signal reduces intractable cookies by 30%, and further rejection of the banner can lead to an additional 32% reduction on subsequent visits, resulting in an average of approximately 11.9 intractable cookies per domain.

5.4 Impact of Website Rank

Figure 7 depicts the average number of intractable cookies set and sent by Tranco websites as per their rank tier. For this analysis, we use the *RandComb* run, which contains 5,915 accepted domains in the stateful phase and 5,121 rejected domains in the stateless phase. The red line shows the average number of intractable cookies sent based on the top list rank of the rejected domains. Accordingly, the

[7] Adobe Audience Manager is a Data Management Platform (DMP) that collects, manages, and segments user data for personalized advertising and audience targeting.

[8] https://globalprivacycontrol.org/orgs

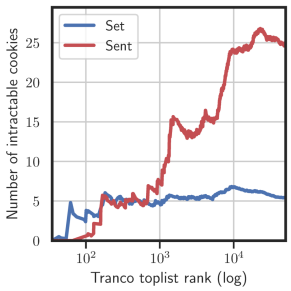
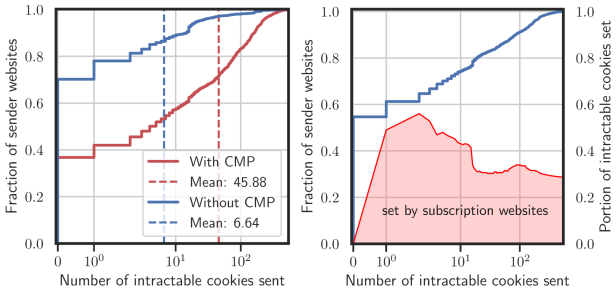


Figure 7: Average number of set and sent intractable cookies over Tranco's toplist ranks.



(a) Impact of CMP **(b) Impact of Cookie Paywall**

Figure 8: Impact of type of banners on the number of intractable cookies for RandComb run.

blue line shows the average number of set cookies over the top list rank of the accepted domains.

For the red line, we observe an ascending trend where popular websites, on average, send fewer intractable cookies than less popular ones. For instance, we observe that the top 50 websites send, on average, zero intractable cookies.[9] Whereas the top 10k websites send, on average, about 25 intractable cookies. This trend may be explained by the tendency of more popular websites to utilize their own resources, potentially resulting in fewer HTTP requests to third parties and, consequently, fewer intractable cookies being sent. Conversely, the blue line remains constant, with websites setting around 5 intractable cookies on average, regardless of their relative ranking.

5.5 Impact of Banner Type

CMP Banners: As mentioned in § 5.2, Consent Management Providers (CMPs) are another entity involved in the context of intractable cookies. Out of 5,121 rejected banners in the *RandComb* run, BannerClick classifies 2,386 as CMP banners. Figure 8a illustrates the difference in the number of intractable cookies between websites with and without CMPs. On average, websites embedding CMP banners set 6.91 times more intractable cookies than those using native banners, highlighting a significant discrepancy in cookie deployment. We also find that CMP banners are relatively harder to reject, as ≈ 40% of them require exploring of "Settings" menu by BannerClick, involving more than one click to be fully rejected. Whereas over 90% of native banners have a direct "Reject" button and can be rejected with a single click. Overall, our findings suggest that CMPs often do not prioritize facilitating user consent or ensuring strict compliance with privacy regulations. They are generally harder to reject and tend to transmit significantly more intractable cookies than native banners.[10] This aligns with the origin of many CMPs in the IAB Europe Transparency and Consent Framework, which was developed by an industry organization representing the interests of the online advertising sector.

Cookie Paywalls: In contrast to typical cookie banners, another form—known as a "cookie paywall"—offers more restricted options to users. Cookie paywalls require users to either opt in to banner policies (mostly tracking) or pay for an ad-free browsing experience through a subscription [64]. Figure 8b displays the ECDF graph for the *RandComb* run, highlighting the portion of intractable cookies set by websites that display cookie paywall-based banners (totaling 90 detected by BannerClick out of 5,915 accepted websites). Note that, following the blue line, 60% of websites send a maximum of 2 intractable cookies, of which around 55% of the cookies are set or reset by websites with cookie paywall banners (peak of the red area in the graph). Interestingly, this proportion drops to around 30% and remains relatively constant for websites that send more than 20 intractable cookies. Overall, this shows that even if users somehow manage to reject all other banners, they still get substantial intractable cookies from a few cookie paywalls.

5.6 Characteristics of Intractable Cookies

Expiration and Renewal Analysis: As mentioned in § 4, we standardize the expiration times of all retrieved cookies to a uniform date far in the future (Saturday, 01 January 2028, 12:12:12 UTC) to increase the consistency of our measurements and enhance the reproducibility of the findings. We recognize that this approach of artificially extending cookie expiration times may introduce bias into our analysis. Additionally, the diversity and the number of cookies users encounter can vary based on their individual browsing behaviors and the websites they visit. Thus, to better assess the validity of our results, we examine the actual expiration times of intractable cookies in the *Cookie Jar*, as well as the number of times they are set or reset across different accepted websites. More detailed analyses and plots are presented in Appendix D.

Figure 9 illustrates the distribution of unique[11] intractable cookies in the popularity runs, based on their expiration times and the number of websites that set them upon banner acceptance. On the x-axis, expiry times are segmented into 1 day, 10 days, 3 months, and 1 year intervals, with "Session" included as a special category for session cookies. These are the common expiry durations for cookies. For example, nearly 25% of intractable cookies have an expiry time set to 1 year (see Appendix D). The y-axis represents buckets corresponding to the percentage of accepted websites setting intractable cookies, including a distinct category for cookies

[9]This observation is supported by our analysis of popularity runs, where none of the 6 rejected domains in the top 50 sent intractable cookies.
[10]These differences may be due to the likelihood that websites with complex ad interdependencies are more inclined to use CMPs and communicate with trackers.
[11]Cookies are grouped by their **name** and **host** attributes.

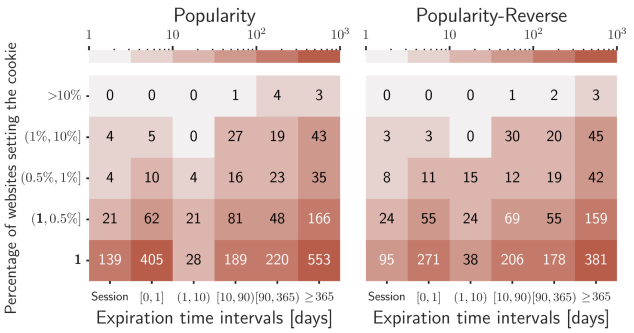


Figure 9: Distribution of intractable cookies based on the number of websites setting them and expiration time. The first row represents a single website setting the cookies.

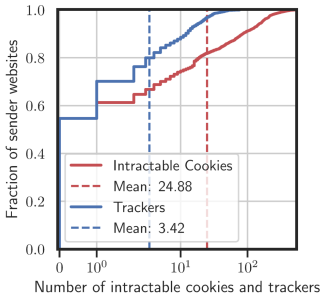


Figure 10: Comparison of the number of intractable cookies and the number of associated trackers per *sender website*.

set only once (i.e., the first row labeled with the bolded 1). The numbers displayed within the heatmap cells represent the count of intractable cookies. For instance, the bottom-right cell of the heatmap for the *Popularity* run shows that 553 unique intractable cookies have an expiry time exceeding one year and were set by a single website during the entire crawl. Both heatmaps show that about 90% of cookies are set by no more than 1% of websites. Furthermore, our analysis reveals that approximately 40% of these cookies have expiration dates of at least one year, indicating a common practice among websites to track users over extended periods. Additionally, over 65% are set with an expiration time of at least 10 days (see Appendix D for details), which aligns with the duration needed to complete the stateful phase and populate the *Cookie Jar* across all four runs. Therefore, for more than 65% of intractable cookies, we can state that our adjustment of expiration times does not lead to an overestimation of their prevalence. However, we cannot draw conclusive insights regarding session cookies or those with shorter expiration durations. A similar distribution pattern is observed in the random runs.

Cookie Synchronization Analysis: Additionally, trackers can employ various techniques—such as cookie synchronization [10, 60]—which enable them to link identifiers across domains and construct more comprehensive user profiles. Following the methodology used in prior works, we examine the transmission of cookie values via redirection URL parameters. We find that, out of 2,545 unique intractable cookies in the *RandComb* run, 76 (i.e., 3%) are synchronized at least once with other *tracker domain* through redirection parameters. Moreover, we manually inspect the values of the top 100 intractable cookies and find that they predominantly contain encrypted or encoded strings, likely serving as unique user identifiers or as part of mechanisms for continued tracking. A smaller portion of cookies contain simple values such as numbers or binary flags (e.g., "YES", "true"), which appear to store preferences or session states. For our cookie syncing detection, we exclude these simpler cases and focus only on cookies containing encrypted or encoded strings longer than 10 characters.

5.7 Domain Analysis

In this section, we analyze the roles and prevalence of *sender website* and *tracker domain* in the transmission of intractable cookies.

The ECDF plots in Figure 10 compare the number of intractable cookies and the number of *tracker domains* (responsible for eliciting them) corresponding to the *sender websites* for the *RandComb* run. On average, each *sender website* is associated with 3.42 different *tracker domains*, each with at least one intractable cookie. Furthermore, based on the means, a single *sender website* has an average of 7.3 intractable cookies per *tracker domain*.

In Table 3, we present the top 20 *tracker domains* along with the total and unique number of intractable cookies associated with them, sorted by the number of *sender websites* responsible for dispatching them. We observe that most top trackers use only a handful of unique cookies to track users across hundreds of the 5,121 rejected websites. We manually verified that the majority of the top 50 domains owning intractable cookies belong to recognized ad tech companies specializing in programmatic advertising solutions (e.g., `pubmatic.com`), while others provide analytic tools and feedback through session recordings and surveys (e.g., `hotjar.com`).

5.8 Partitioned Cookies Analysis

Partitioned cookies are a newly proposed privacy feature in Chrome [31], also adopted by browsers like Firefox. They scope third-party cookies to the top-level site, preventing them from being shared across different websites. Accordingly, they have the potential to mitigate intractable cookies by isolating their transmission to each *setter website*. For example, in the scenario described in § 5.2, when `basic.com` sets a cookie from `tracker.net`, the cookie is stored in a partitioned context specific to `basic.com`. Later, when the user visits `new.com` (*sender website*), even if it loads a resource from `tracker.net`, the browser does not send the previously set cookie ($id = 123$) because partitioned storage treats `new.com` as a separate context. As a result, `tracker.net` cannot correlate the user's activity across `basic.com` and `new.com`.

As detailed in § 4.1, we conduct a separate run using Chrome to measure the prevalence of partitioned cookies. The results reveal a distribution of cookie types similar to the other runs conducted with Firefox. As shown in Table 4, out of 79,898 unique stored cookies, 521 are marked partitioned. However, among 3,177 unique tracking cookies, only 40 (1.3%) are partitioned, of which 26 are accompanied by non-partitioned tracking cookies from the same tracker, with 9 having the same value.[12] Interestingly, we also observe that Chrome

[12]In most cases, the partitioned cookie has a similar `name`, appended by 'p' or '_p'.

	Tracker Domain	#Cookies	#UniqCookies	#Senders		Tracker Domain	#Cookies	#UniqCookies	#Senders
1	amazon-adsystem.com	1,526	2	763	11	3lift.com	1,240	5	365
2	adsrvr.org	1,344	2	672	12	lijit.com	1,084	13	343
3	criteo.com	2,623	7	649	13	bidswitch.net	2,723	11	323
4	pubmatic.com	19,905	40	621	14	a-mo.net	5,458	18	311
5	adnxs.com	3,020	5	604	15	taboola.com	6,118	42	293
6	casalemedia.com	2,128	4	532	16	nr-data.net	282	1	282
7	id5-sync.com	1,054	7	517	17	bidr.io	414	2	282
8	openx.net	1,203	3	401	18	tapad.com	771	3	257
9	smartadserver.com	2,074	7	375	19	liadm.com	444	2	243
10	sharethrough.com	2,357	12	366	20	quantserve.com	454	2	227

Table 3: Top 20 *tracker domains* with the total and unique number of intractable cookies associated with them sorted by the *sender websites* sending them.

Cookie Type	Total Unique	Partitioned	Along with NP
All Cookies	79,898	521	-
Tracking Cookies	3,177	40	26

Table 4: Summary of unique partitioned cookies and those accompanying the non-partitioned (NP) cookies from the same *tracker domain*.

does not overwrite existing cookies when their partitioned attribute differs. In other words, the same cookie can be set twice—once with the partitioned attribute and once without—both of which are sent in subsequent HTTP requests.

Overall, despite nearly three years since Google introduced partitioned cookies through the Cookies Having Independent Partitioned State (CHIPS) initiative with the release of Chrome 100 in March 2022, their adoption remains limited and gradual. This finding aligns with a recent study on CHIPS [79]. Nevertheless, a longitudinal study is needed to assess whether partitioned cookies will achieve widespread adoption and be effectively applied in real-world scenarios. We discuss partitioned cookies further in § 6.

6 DISCUSSION

In this section, we examine the relationship between banner designs, intractable cookies, and privacy regulations, exploring mitigation strategies such as partitioned cookies and their limitations. We then introduce a Browser-Integrated Consent Mechanism as our proposed approach to addressing the shortcomings of the current consent mechanism that lead to intractable cookies.

6.1 Interpretation of Privacy Regulations

The existence of intractable cookies can be traced back to the impractical implementation of consent mechanisms, primarily in the form of cookie banners, introduced in response to the ePrivacy Directive and shaped further by the consent requirements defined under the GDPR. In the following, we examine the current deployment of cookie banners by addressing two key limitations in the broader privacy framework: the fragmented interpretation of the valid consent requirements and the ambiguity regarding which entities are responsible for ensuring compliance.

Fragmented Interpretation of Consent Requirements: While the GDPR outlines core criteria for valid consent, such as being freely given, specific, informed, and unambiguous, it does not prescribe how these requirements should be technically implemented. Instead, practical interpretation has been shaped through soft law instruments, including guidelines from the European Data Protection Board (EDPB) [25] and national Data Protection Authorities (DPAs) [11],[13] as well as court rulings.[14]

Although the EDPB uses the GDPR as its legal basis to harmonize data protection practices across the EU and to interpret vague or open-textured provisions, individual DPAs retain discretion in how these interpretations are applied and enforced at the national level. This can lead to divergent regulatory outcomes among Member States. For instance, the EDPB's Cookie Banner Taskforce Report (2023) [27] notes that most DPAs consider the absence of a "Reject" option on any layer where a consent ("Accept") button is present to be non-compliant with the ePrivacy Directive. However, some authorities, such as the Spanish Data Protection Authority (AEPD), have adopted a more permissive stance, allowing the "Reject All" option to appear only in a secondary layer. This regulatory divergence contributes to a fragmented enforcement landscape. In practice, recent studies report widespread non-compliance and inconsistent implementation of cookie banners across websites [52, 67].

Subsequently, as mentioned in § 4.3, many cookie banners place the "Reject All" option within "Settings" layers, requiring users to navigate through granular settings often spread across multiple tabs. This design introduces unnecessary friction and undermines the effectiveness of cookie banners in clearly communicating data collection practices [42, 53, 76, 77]. Even when granular choices are necessary, they could be presented in a more user-friendly manner—accessible yet unobtrusive to the majority of users.

Moreover, although the GDPR provides a general definition of valid consent, terms such as "freely given" and "unambiguous" remain open to interpretation. Prior research shows that cookie banners frequently employ deceptive patterns that nudge users toward acceptance through habituation rather than informed

[13]The European Data Protection Board (EDPB) is an independent EU body that ensures consistent application of the GDPR across member states. On the other hand, National Data Protection Authorities (DPAs) are country-level regulators responsible for enforcing data protection laws and issuing context-specific guidance.
[14]For example, the ECJ's *Planet49* decision explicitly ruled that pre-ticked checkboxes do not constitute valid consent [1].

choice [9, 17, 59, 71]. The rise of cookie paywalls [64] further pressures users into accepting tracking, often against their preferences. As a result, many users reluctantly consent to tracking on certain sites, with little to no ability to revoke that consent later [53]. As we showed, many of these seemingly isolated acceptances propagate across visits via intractable cookies, enabling inter-domain tracking even before users provide explicit consent on the visited site.

Ambiguity of Accountability: The transmission of intractable cookies to *tracker domains* raises concerns under existing privacy regulations. The GDPR requires valid user consent before collecting or processing personal data, unless another legal basis applies—such as necessity for a service explicitly requested by the user (Article 6). While the GDPR designates the "data controller"[15] as responsible for ensuring compliance (Article 24), accountability becomes complex when multiple parties are involved. Article 26 introduces the concept of "joint controllers" and requires an *arrangement* to define their respective responsibilities.

According to the EDPB Guidelines [26], joint controllership arises when two or more entities jointly determine the purposes and means of data processing, regardless of whether they have equal access to the data. This was affirmed by the CJEU in the *Fashion ID* ruling,[16] establishing that shared decision-making alone can trigger shared responsibility. In practice, however, applying this principle is challenging within the web's complex tracking ecosystems, involving multiple actors such as websites, third-party trackers, advertisers, DMPs, and CMPs. These actors rarely establish or disclose clear joint controller arrangements, making it difficult to determine who is accountable for informing users or fulfilling their data rights. This lack of clarity hinders enforcement and risks leaving users without a clear path to contest or revoke consent, ultimately limiting the GDPR's effectiveness in protecting personal data in such distributed environments.

In a nutshell, the current deployment of cookie banners often reflects ad hoc compliance efforts, aimed more at fulfilling legal formalities than enabling meaningful user choice, creating a potential *false sense of privacy*. While privacy regulations promote transparency and accountability, their effectiveness hinges on a clear understanding of user behavior and the roles of various actors in the consent ecosystem. Without coordinated input from regulators, developers, and interdisciplinary experts, these frameworks risk undermining the very privacy they seek to protect by inconsistent enforcement and superficial implementation.

6.2 Intractable Cookies Mitigation

As discussed in § 5.2, from a technical perspective, the core cause of intractable cookies may be the *sender websites*' inability to determine the existence of cookies set by previously visited websites (i.e., *setter websites*). Consequently, several solutions can mitigate or eliminate the intractable cookies phenomenon. One approach is to *prevent the loading of third-party resources* before banner acceptance. However, given the current structure of consent mechanisms—where websites handle user preferences via banners and browsers control request transmission—this solution is not feasible.

Alternatively *blocking third-party cookies* entirely is the most straightforward approach to mitigate the privacy-intrusive nature of tracking cookies, including intractable cookies. This approach has already been implemented by browsers like Safari as a default setting [29]. As for Chrome, which accounts for over 65% of the browser market share across both desktop and mobile platforms [2], privacy-conscious users have the option to customize their settings and block third-party cookies. However, studies [53] show that most users are unaware of these controls or the tracking technologies behind them. More importantly, the debate over tracking extends beyond individual user preferences, as it involves conflicting interests among users, publishers, and advertisers. These competing priorities complicate the feasibility of outright blocking. For instance, Google's July 2024 reversal [63] of its 2020 pledge [16, 74] to phase out third-party cookies in Chrome underscores the tension between privacy advocacy and economic interests.

Nonetheless, if we focus solely on users' interests and assume that blocking third-party cookies enhances their web experience by improving user privacy, the reality is more complex. Entities dependent on advertising revenue will likely shift to alternative tracking methods, such as fingerprinting [4, 41], or adjust their pricing strategies to compensate for the loss of targeted ads [54]. Consequently, from a broader perspective, eliminating third-party cookies entirely may not provide the anticipated benefits for users unless an alternative monetization model is introduced.

Another possible solution to mitigate intractable cookies is *partitioned cookies*. However, besides its limited implementation (see § 5.8), partitioned cookies also have several other limitations:

- *Developer Reliance:* Approaches that depend on widespread developer adoption often fail to achieve satisfactory results. For instance, studies on Content Security Policy (CSP) [46, 66] reveal that less than 2% of websites correctly implement it, with most deployments being ineffective or poorly configured. Although implementing partitioned cookies is less complex than CSP, it still requires modifying attributes like `SameSite` and appending the `_Host` prefix to handle subdomains.
- *Lack of Incentives:* The incentive for adopting partitioned cookies remains unclear (particularly when considering the *tracker domain* as the owner of the cookies) unless explicitly enforced by regulations. In comparison, CSP adoption is driven by its direct relation to the website's own security.
- *Incompatibility with Consent-Based Tracking:* Most importantly, partitioned cookies lack the technical capability to enable inter-domain tracking, even upon explicit users' consent. This limitation undermines consent-based tracking, as it prevents users from selectively allowing or blocking tracking based on their preferences via banners. For instance, some users may wish to receive personalized advertising while blocking tracking from specific websites, such as those handling sensitive content. However, with partitioned cookies, such granular control is no longer possible, rendering cookie banners ineffective for managing tracking preferences.

[15] Article 4(7) GDPR defines the *data controller* as "the natural or legal person, public authority, agency, or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".
[16] In *Fashion ID* (C-40/17), the Court held that a website embedding a third-party plugin (e.g., Facebook's Like button) could be a joint controller, even without accessing the collected data, if it contributes to the determination of purposes and means.

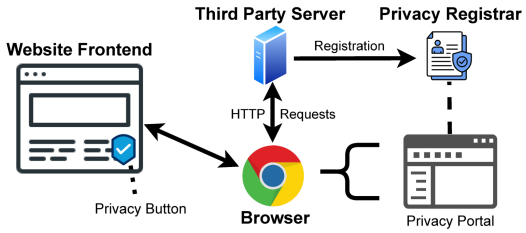


Figure 11: Overview of the proposed Browser-Integrated Consent Mechanism, consisting of four major components: browser, website, third party, and registrar.

6.3 Browser-Integrated Consent Mechanism

As discussed in § 6.1, one of the major drawbacks of the GDPR is its ambiguity in defining the "data controller" as the entity responsible for collecting and handling user consent. Moreover, obtaining user consent is distinct from data collection and processing and could instead be managed by a separate entity within the data flow. Accordingly, given the current structure of the web ecosystem—where the browser serves as the central element orchestrating communication between various entities—a practical approach is to integrate the consent mechanism directly into the browser. This would streamline the consent process, ensuring a more consistent and efficient implementation of consent mechanisms that can be easily scaled and updated.

In this model, browsers serve as the intermediary entity responsible for collecting user consent and preferences and applying them accordingly. This approach reduces the burden on both users and publishers (i.e., developers). Users can configure their preferences through a well-structured consent management portal within the browser, ensuring their choices are consistently applied across all websites they visit. Likewise, publishers would no longer need to implement their own consent banners, resulting in a more uniform and user-friendly design.

Figure 11 depicts an overview of our proposed browser-integrated consent mechanism, which consists of four main components: the browser, the website, third parties (e.g., trackers or advertisers), and a privacy registrar. In this scheme, each third party intending to engage in cross-site tracking (e.g., via cookies) must register with a centralized registrar and disclose its intended purposes, similar in spirit to how vendors register with the IAB Europe Global Vendor List in the context of CMPs [38]. However, unlike the IAB, which primarily represents the online advertising industry, the registrar in our proposal requires a more neutral and trustworthy governance model. Possible implementations include a government-backed registrar or a multi-stakeholder consortium similar to those organized by W3C [78]. The browser retrieves declarations from this registrar and uses them to decide, based on the user's privacy preferences, whether to allow or block specific third-party requests.

Users interact with consent controls through a unified, browser-operated system called the *privacy portal*, which offers a consistent interface for managing preferences per website and registered third parties. The portal can be accessed through two entry points: (1) directly via the browser's settings, and (2) through a *privacy button* displayed in the corner of the viewport when visiting a website, which opens the relevant section of the portal for that site. For example, users can opt out of specific trackers or websites based on category, or adjust their settings at any time using the privacy button. The interface's structure and appearance remain consistent across all websites. Within the portal, visible "Accept All" and "Reject All" buttons are always available, alongside more granular controls populated using data retrieved from the privacy registrar.

This high-level prototype can be further developed and refined to strengthen user privacy while maintaining essential advertising capabilities. One potential approach is for browsers to treat all third-party cookies as partitioned by default and transmit them cross-site only if explicitly permitted by user preferences. Furthermore, the portal could incorporate a subscription-based model, allowing websites to monetize their content directly as an alternative to ad-supported tracking, similar to existing cookie paywalls. In this model, websites could adapt their behavior and render the front end based on user choices or subscription status. Although this model may create a divide between paying and non-paying users, its alignment with free market principles could, in the long run, lead to improved services and a more sustainable online environment.

Lastly, the browser-integrated approach aims to reduce complexity by consolidating consent interactions and preference management into a single, user-centric system. Unlike prior browser-level mechanisms such as Global Privacy Control (GPC) and Advanced Data Protection Control (ADPC) [37], our proposal operates independently of publisher support, as browsers can now allow or block third-party requests based on the user's privacy preferences set via the privacy portal and information retrieved from the registrar. This independence makes it more practical and enforceable in real-world scenarios, particularly when dealing with non-compliant or uncooperative websites. The primary challenge, however, lies in persuading or requiring browser vendors to adopt such a mechanism, an objective that could be facilitated through regulatory mandates or privacy-focused legislation.

7 CONCLUSION

In this paper, we reveal the prevalence of *intractable* cookies—tracking cookies that are set by websites where users accept their banners, persistent in the user browser, and sent to tracking domains before the user's explicit consent on subsequent websites. Through extensive measurements involving 20,000 domains from the Tranco top list, we demonstrated that around 50% of the websites sent at least one intractable cookie. Furthermore, we assessed how banner interaction, enabling GPC signal, can contribute to preventing intractable cookies. We then explored the impact of the website rank and type of the banner on the prevalence of these cookies. Moreover, we analyzed the expiration and renewal characteristics of intractable cookies, along with their domain distribution. Finally, we examined current technical solutions, such as partitioned cookies, that aim to mitigate intractable cookies and discussed their limitations.

Overall, our findings highlight a gap between the technical and legislative aspects of the web tracking ecosystem, leading to solutions like cookie banners that add complexity without effectively addressing the core issue. We advocate for meaningful improvements through the conscious collaboration of all stakeholders, including developers, regulators, publishers, and advertisers.

Acknowledgments

We sincerely thank the anonymous reviewers for their thoughtful feedback, which significantly improved the readability and clarity of the paper, and acknowledge the Max Planck Institute for Informatics for funding the work and providing the infrastructure needed for the measurements and analysis.

References

[1] 2019. Judgment of the Court (Grand Chamber), Planet49 GmbH v Bundesverband der Verbraucherzentralen und Verbraucherverbände — Verbraucherzentrale Bundesverband eV. https://curia.europa.eu/juris/document/document.jsf?docid=218462&doclang=EN. Accessed: 2025-05-13.

[2] 2025. Browser Market Share Worldwide - January 2025. https://gs.statcounter.com/browser-market-share. Accessed: 2025-02-25.

[3] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. 2014. The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. Association for Computing Machinery, New York, NY, USA, 674–689. https://doi.org/10.1145/2660267.2660347

[4] Alex Angove-Plumb. 2022. Browser Fingerprinting and the Death of Third-Party Cookies. *CHOICE* (2022). https://doi.org/10.13140/RG.2.2.45678.90123

[5] France Belanger, Janine S Hiller, and Wanda J Smith. 2002. Trustworthiness in Electronic Commerce: The Role of Privacy, Security, and Site Attributes. *The Journal of Strategic Information Systems* 11, 3-4 (2002), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

[6] Tiago Bianchi. 2025. Google: advertising revenue 2024. https://www.statista.com/statistics/266249/advertising-revenue-of-google/ Accessed: 2025-02-23.

[7] Nataliia Bielova. 2017. Web Tracking Technologies and Protection Mechanisms. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17)*. Association for Computing Machinery, New York, NY, USA, 2607–2609. https://doi.org/10.1145/3133956.3136067

[8] Nataliia Bielova, Laura Litvine, Anysia Nguyen, Mariam Chammat, Vincent Toubiana, and Estelle Harry. 2024. The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions. In *33rd USENIX Security Symposium (USENIX Security 24)*. USENIX Association, Anaheim, CA. https://doi.org/10.5555/3691330.3691356

[9] Ida Borberg, René Hougaard, Willard Rafnsson, and Oksana Kulyk. 2022. "So I Sold My Soul": Effects of Dark Patterns in Cookie Notices on End-User Behavior and Perceptions. In *Proceedings of the 2022 Usable Security and Privacy Symposium (USEC '22)*. Internet Society, San Diego, USA, 1–11. https://doi.org/10.14722/usec.2022.23001

[10] Ahmed Bouhoula, Karel Kubicek, Amit Zac, Carlos Cotrini, and David A. Basin. 2024. Automated Large-Scale Analysis of Cookie Notice Compliance. In *33rd USENIX Security Symposium (USENIX Security 24)*. USENIX Association, Philadelphia, PA, 1723–1739. https://doi.org/10.5555/3691330.3691345

[11] Bundesbeauftragter für den Datenschutz und die Informationsfreiheit (BfDI). 2023. Cookie-Banner. https://www.bfdi.bund.de/DE/Buerger/Inhalte/Telemedien/Cookie-Banner.html. Accessed: 2025-05-06.

[12] California Office of the Attorney General. 2021. California Consumer Privacy Act (CCPA) - Global Privacy Control. https://oag.ca.gov/privacy/ccpa/gpc Accessed: 2025-02-28.

[13] Ed Chau and Robert Hertzberg. 2018. California Consumer Privacy Act. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375. Accessed: 2025-05-13.

[14] Quan Chen, Panagiotis Ilia, Michalis Polychronakis, and Alexandros Kapravelos. 2021. Cookie Swap Party: Abusing First-Party Cookies for Web Tracking. In *Proceedings of the 2021 World Wide Web Conference (WWW '21)*. Association for Computing Machinery, New York, NY, USA, 2117–2129. https://doi.org/10.1145/3442381.3449837

[15] European Commission. 2018. The General Data Protection Regulation (GDPR). https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Accessed: 2024-05-01.

[16] Cookiebot. 2022. Google ending third-party cookies in Chrome. https://www.cookiebot.com/en/google-third-party-cookies/. Accessed: 2025-05-13.

[17] Lynne M Coventry, Debora Jeske, John M Blythe, James Turland, and Pam Briggs. 2016. Personality and Social Framing in Privacy Decision-Making: A Study on Cookie Acceptance. *Frontiers in Psychology* 7 (2016), 1341. https://doi.org/10.3389/fpsyg.2016.01341

[18] Adrian Dabrowski, Georg Merzdovnik, Johanna Ullrich, Gerald Sendera, and Edgar Weippl. 2019. Measuring Cookies and Web Privacy in a Post-GDPR World. In *Passive and Active Measurement (PAM)*. Springer International Publishing, Cham, 258–270. https://doi.org/10.1007/978-3-030-15986-3_17

[19] Martin Degeling, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz. 2019. We Value Your Privacy... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. In *Proceedings of the 2019 Network and Distributed System Security Symposium (NDSS '19)*. Internet Society, San Diego, USA, 345–346. https://doi.org/10.23919/NDSS.2019.8712210

[20] Nurullah Demir, Jan Hörnemann, Matteo Große-Kampmann, Tobias Urban, Norbert Pohlmann, Thorsten Holz, and Christian Wressnegger. 2023. On the Similarity of Web Measurements Under Different Experimental Setups. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23)*. Association for Computing Machinery, New York, NY, USA, 356–369. https://doi.org/10.1145/3618257.3624795

[21] Statista Research Department. 2025. Global Meta advertising revenue 2024. https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/ Accessed: 2025-02-23.

[22] David Dittrich, Erin Kenneally, Michael Bailey, and Dennis Fischer. 2012. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. *U.S. Department of Homeland Security* 7, 3 (2012), 1–22. https://doi.org/10.2139/ssrn.2445102

[23] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. 2013. ZMap: Fast Internet-wide Scanning and Its Security Applications. In *Proceedings of the 22nd USENIX Security Symposium (USENIX Security 13)*. USENIX Association, Washington, D.C., 605–620. https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric

[24] Steven Englehardt and Arvind Narayanan. 2016. Online Tracking: A 1-million-site Measurement and Analysis. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*. Association for Computing Machinery, New York, NY, USA, 1388–1401. https://doi.org/10.1145/2976749.2978313

[25] European Data Protection Board. 2018. About the EDPB. https://edpb.europa.eu/about-edpb/about-edpb_en. Accessed: 2025-05-06.

[26] European Data Protection Board. 2021. Guidelines 07/2020 on the concepts of controller and processor in the GDPR (Version 2.1). https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_en. Accessed: 2025-05-06.

[27] European Data Protection Board. 2023. Report on the work undertaken by the Cookie Banner Taskforce. https://edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-cookie-banner-taskforce_en. Accessed: 2025-05-06.

[28] European Parliament and the Council. 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). https://eur-lex.europa.eu/eli/dir/2002/58/oj/eng. Accessed: 2025-05-06.

[29] Jon Fingas. 2020. Safari Now Blocks All Third-Party Cookies by Default. *Engadget* (2020). https://doi.org/10.13140/RG.2.2.12345.67890

[30] Imane Fouad, Cristiana Santos, Arnaud Legout, and Nataliia Bielova. 2022. My Cookie is a Phoenix: Detection, Measurement, and Lawfulness of Cookie Respawning with Browser Fingerprinting. *Proceedings on Privacy Enhancing Technologies* 2022, 3 (2022), 79–98. https://doi.org/10.56553/POPETS-2022-0063

[31] Google. 2024. Cookies Having Independent Partitioned State (CHIPS) - Privacy Sandbox. https://developers.google.com/privacy-sandbox/3pcd/chips. Accessed: 2024-05-15.

[32] Google. 2025. How Google uses cookies for advertising. https://business.safety.google/adscookies/?utm_source=chatgpt.com Accessed: 2025-02-27.

[33] Global Privacy Control Group. 2024. Global Privacy Control (GPC). https://privacycg.github.io/gpc-spec/. Accessed: 2024-10-14.

[34] Ralf Gundelach and Dominik Herrmann. 2023. CookieScanner: An Automated Tool for Detecting and Evaluating GDPR Consent Notices on Websites. In *Proceedings of the 18th International Conference on Availability, Reliability and Security (ARES '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3600160.3605000

[35] Matthias Götze, Srdjan Matic, Costas Iordanou, Georgios Smaragdakis, and Nikolaos Laoutaris. 2022. Measuring Web Cookies in Governmental Websites. In *Proceedings of the 14th ACM Web Science Conference (WebSci '22)*. Association for Computing Machinery, Barcelona, Spain, 44–54. https://doi.org/10.1145/3501247.3531563

[36] Maximilian Hils, Daniel W Woods, and Rainer Böhme. 2020. Measuring the Emergence of Consent Management on the Web. In *Proceedings of the 2020 ACM Internet Measurement Conference (IMC '20)*. 317–332. https://doi.org/10.1145/3419394.3423622

[37] Soheil Human. 2022. Advanced Data Protection Control (ADPC): An Interdisciplinary Overview. *arXiv preprint arXiv:2209.09724* (2022). https://arxiv.org/abs/2209.09724

[38] IAB Europe. 2022. What is TCF v2.0? https://iabeurope.eu/tcf-2-0/. Accessed: 2025-05-13.

[39] IAB Europe. 2022. What is the Transparency & Consent Framework (TCF)? https://iabeurope.eu/transparency-consent-framework/. Accessed: 2025-05-13.

[40] Umar Iqbal, Pouneh Nikkhah Bahrami, Rahmadi Trimananda, Hao Cui, Alexander Gamero-Garrido, Daniel J. Dubois, David Choffnes, Athina Markopoulou, Franziska Roesner, and Zubair Shafiq. 2023. Tracking, Profiling, and Ad Targeting

in the Alexa Echo Smart Speaker Ecosystem. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23)*. Association for Computing Machinery, New York, NY, USA, 569–583. https://doi.org/10.1145/3618257.3624825

[41] Umar Iqbal, Steven Englehardt, and Zubair Shafiq. 2021. Fingerprinting the Fingerprinters: Learning to Detect Browser Fingerprinting Behaviors. In *Proceedings of the 2021 IEEE Symposium on Security and Privacy (SP '21)*. 1231–1246. https://doi.org/10.1109/SP40001.2021.00097

[42] Nikhil Jha, Martino Trevisan, Marco Mellia, Rodrigo Irarrazaval, and Daniel Fernandez. 2023. I Refuse if You Let Me: Studying User Behavior with Privacy Banners at Scale. In *Proceedings of the 2023 Network Traffic Measurement and Analysis Conference (TMA '23)*. IEEE, Napoli, Italy, 1–9. https://doi.org/10.1109/TMA59887.2023.10304876

[43] Nikhil Jha, Martino Trevisan, Luca Vassio, and Marco Mellia. 2022. The Internet with Privacy Policies: Measuring the Web Upon Consent. *ACM Transactions on the Web* 16, 3 (2022), 1–24. https://doi.org/10.1145/3548684

[44] Justdomains. 2022. DOMAIN-ONLY Filter Lists. https://github.com/justdomains/blocklists. Accessed: 2025-05-13.

[45] JustDomains. 2025. Automated scripts to support converting filter lists to "domain-only" lists. https://github.com/justdomains/ci Accessed: 2025-02-27.

[46] Christoph Kerschbaumer, Sid Stamm, and Stefan Brunthaler. 2016. Injecting CSP for Fun and Security. In *Proceedings of the 2016 International Conference on Information Systems Security and Privacy (ICISSP '16)*. https://doi.org/10.5220/0005716201230134

[47] 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. https://doi.org/10.1093/ijlit/eat008

[48] Michael Kretschmer, Jan Pennekamp, and Klaus Wehrle. 2021. Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web. *ACM Transactions on the Web* 15, 4 (2021). https://doi.org/10.1145/3466722

[49] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. https://tranco-list.eu/. Accessed: 2025-05-13.

[50] Chung Hun Lee and David A Cranage. 2011. Personalisation–Privacy Paradox: The Effects of Personalisation and Privacy Assurance on Customer Responses to Travel Web Sites. *Tourism Management* 32, 5 (2011), 987–994. https://doi.org/10.1016/j.tourman.2010.08.011

[51] Dominique Machuletz and Rainer Böhme. 2020. Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs After GDPR. *Proceedings on Privacy Enhancing Technologies* 2 (2020), 481–498. https://doi.org/10.2478/popets-2020-0039

[52] Célestin Matte, Nataliia Bielova, and Cristiana Santos. 2020. Do Cookie Banners Respect My Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. In *Proceedings of the 2020 IEEE Symposium on Security and Privacy (SP '20)*. IEEE, San Francisco, USA, 791–809. https://doi.org/10.1109/SP40000.2020.00076

[53] Maryam Mehrnezhad, Kovila Coopamootoo, and Ehsan Toreini. 2021. How Can and Would People Protect from Online Tracking? *Proceedings on Privacy Enhancing Technologies* 1 (2021), 105–125. https://doi.org/10.2478/popets-2021-0007

[54] Klaus M. Miller and Bernd Skiera. 2024. Economic consequences of online tracking restrictions: Evidence from cookies. *International Journal of Research in Marketing* 41, 2 (2024), 241–264. https://doi.org/10.1016/j.ijresmar.2023.10.001

[55] Mozilla. 2005. Public Suffix List. https://publicsuffix.org/. Accessed: 2025-05-13.

[56] Mozilla. 2021. Total Cookie Protection in Firefox. https://blog.mozilla.org/security/2021/02/23/total-cookie-protection/. Accessed: 2025-05-13.

[57] Shaoor Munir, Sandra Siby, Umar Iqbal, Steven Englehardt, Zubair Shafiq, and Carmela Troncoso. 2023. CookieGraph: Understanding and Detecting First-Party Tracking Cookies. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23)*. Association for Computing Machinery, Copenhagen, Denmark, 3490–3504. https://doi.org/10.1145/3576915.3616586

[58] Midas Nouwens, Rolf Bagge, Janus Bager Kristensen, and Clemens Nylandsted Klokmose. 2022. Consent-O-Matic: Automatically Answering Consent Pop-ups Using Adversarial Interoperability. In *Proceedings of the 2022 ACM CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '22)*. Association for Computing Machinery, 238. https://doi.org/10.1145/3491101.3519683

[59] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark Patterns After the GDPR: Scraping Consent Pop-ups and Demonstrating Their Influence. In *Proceedings of the 2020 ACM CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376321

[60] Emmanouil Papadogiannakis, Panagiotis Papadopoulos, Nicolas Kourtellis, and Evangelos P. Markatos. 2021. User Tracking in the Post-cookie Era: How Websites Bypass GDPR Consent to Track Users. In *Proceedings of the 2021 World Wide Web Conference (WWW '21)*. Association for Computing Machinery, New York, NY, USA, 2130–2141. https://doi.org/10.1145/3442381.3450056

[61] Panagiotis Papadopoulos, Nicolas Kourtellis, and Evangelos Markatos. 2019. Cookie Synchronization: Everything You Always Wanted to Know But Were Afraid to Ask. In *Proceedings of the 2019 World Wide Web Conference (WWW '19)*. Association for Computing Machinery, New York, NY, USA, 1432–1442. https://doi.org/10.1145/3308558.3313542

[62] Craig Partridge and Mark Allman. 2016. Ethical Considerations in Network Measurement Papers. *Commun. ACM* 59, 10 (2016), 58–64. https://doi.org/10.1145/2896816

[63] Aimee Picchi. 2024. Google Reneges on Plan to Remove Third-Party Cookies in Chrome. *CBS News* (2024). https://doi.org/10.13140/RG.2.2.23456.78901

[64] Ali Rasaii, Devashish Gosain, and Oliver Gasser. 2023. Thou Shalt Not Reject: Analyzing Accept-Or-Pay Cookie Banners on the Web. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23)*. Association for Computing Machinery, New York, NY, USA, 154–161. https://doi.org/10.1145/3618257.3624846

[65] Ali Rasaii, Shivani Singh, Devashish Gosain, and Oliver Gasser. 2023. Exploring the Cookieverse: A Multi-Perspective Analysis of Web Cookies. In *Passive and Active Measurement (PAM)*. Springer Nature Switzerland, Cham, 623–651. https://doi.org/10.1007/978-3-031-28486-1_26

[66] Sebastian Roth, Timothy Barron, Stefano Calzavara, Nick Nikiforakis, and Ben Stock. 2020. Complex Security Policy? A Longitudinal Analysis of Deployed Content Security Policies. In *Proceedings of the 2020 Network and Distributed System Security Symposium (NDSS '20)*. https://eref.uni-bayreuth.de/id/eprint/91468/

[67] Cristiana Santos, Nataliia Bielova, and Célestin Matte. 2020. Are Cookie Banners Indeed Compliant with the Law? Deciphering EU Legal Requirements on Consent and Technical Means to Verify Compliance of Cookie Banners. *Technology and Regulation* 2020 (2020), 91–135. https://doi.org/10.26116/techreg.2020.009

[68] Cristiana Santos, Arianna Rossi, Lorena Sanchez Chamorro, Kerstin Bongard-Blanchy, and Ruba Abu-Salma. 2021. Cookie Banners, What's the Purpose? Analyzing Cookie Banner Text Through a Legal Lens. In *Proceedings of the 2021 ACM Workshop on Privacy in the Electronic Society (WPES '21)*. Association for Computing Machinery, New York, NY, USA, 187–194. https://doi.org/10.1145/3463676.3485611

[69] H. Jeff Smith, Sandra J. Milberg, and Sandra J. Burke. 1996. State of the Information Privacy Literature: Where Are We Now and Where Should We Go? *MIS Quarterly* 20, 2 (1996), 167–196. https://doi.org/10.2307/249477

[70] Michael Smith, Antonio Torres-Agüero, Riley Grossman, Pritam Sen, Yi Chen, and Cristian Borcea. 2024. A Study of GDPR Compliance Under the Transparency and Consent Framework. In *Proceedings of the 2024 ACM Web Conference (WWW '24)*. Association for Computing Machinery, Singapore, 1227–1236. https://doi.org/10.1145/3589334.3645618

[71] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. 2020. Circumvention by Design: Dark Patterns in Cookie Consent for Online News Outlets. In *Proceedings of the 2020 ACM Nordic Conference on Human-Computer Interaction (NordiCHI '20)*. Association for Computing Machinery, Tallinn, Estonia, 1–12. https://doi.org/10.1145/3419249.3420132

[72] Jannick Sørensen and Sokol Kosta. 2019. Before and After GDPR: The Changes in Third Party Presence at Public and Private European Websites. In *Proceedings of the 2019 World Wide Web Conference (WWW '19)*. Association for Computing Machinery, San Francisco, USA, 1590–1600. https://doi.org/10.1145/3308558.3313554

[73] 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. https://doi.org/10.2478/popets-2019-0023

[74] Bowdeya Tweh and Sahil Patel. 2020. Google Chrome to Phase Out Third-Party Cookies in Effort to Boost Privacy. *The Wall Street Journal* (2020). https://doi.org/10.13140/RG.2.2.34567.89012

[75] Tobias Urban, Martin Degeling, Thorsten Holz, and Norbert Pohlmann. 2020. Beyond the Front Page: Measuring Third Party Dynamics in the Field. In *Proceedings of the 2020 World Wide Web Conference (WWW '20)*. Association for Computing Machinery, 1275–1286. https://doi.org/10.1145/3366423.3380203

[76] 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)*. Association for Computing Machinery, London, UK, 973–990. https://doi.org/10.1145/3319535.3354212

[77] Nikolas Wehner, Michael Seufert, Raimund Schatz, and Tobias Hoßfeld. 2023. Do You Agree? Contrasting Google's Core Web Vitals and the Impact of Cookie Consent Banners with Actual Web QoE. *Quality and User Experience* 8, 1 (2023), 5. https://doi.org/10.1007/s41233-023-00060-9

[78] World Wide Web Consortium (W3C). 2022. W3C Governance: Overview and Structure. https://www.w3.org/2022/10/bod-governance.html. Accessed: 2025-05-06.

[79] Maximilian Zöllner, Anja Feldmann, and Ha Dao. 2025. A First Look at Cookies Having Independent Partitioned State. In *Passive and Active Measurement (PAM)*. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-55528-2_1

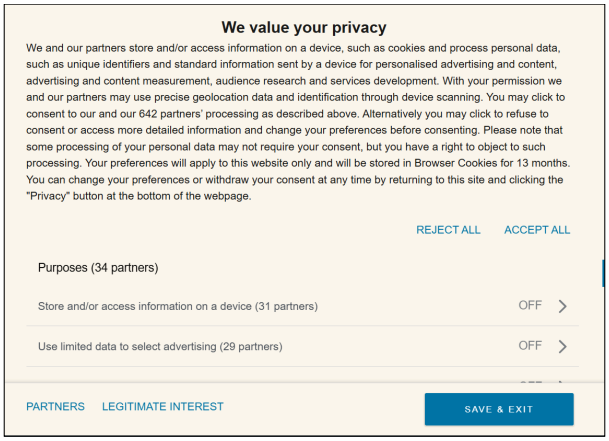


Figure 12: Example of a banner displaying granular options after clicking the "Settings" button. Clicking "SAVE & EXIT" should be equivalent to rejecting tracking, as all non-essential options are disabled by default.

A Ethical Considerations

In conducting our measurements, we abide by the ethical guidelines proposed by Partridge and Allman [62] and Kenneally and Dittrich [22], and follow the best measurement practices as described by Durumeric et al. [23]. Our methodology involves running OpenWPM in an automated manner to crawl each website using a standard web browser configuration. We utilize dedicated measurement machines configured with informative reverse DNS (rDNS) names, which are subdomains of a research-affiliated organizational domain and are easily identifiable. Moreover, we offer stakeholders the option to opt out and be excluded from our measurements. Throughout our measurement period, we did not receive any complaints.

B Banner Screenshot

Figure 12 shows an example of a cookie banner that provides granular choices for different partners and purposes after the user clicks the "Settings" button. In this case, all non-essential cookies are turned off by default. As a result, clicking the "SAVE & EXIT" button should functionally correspond to rejecting the banner.

C Cookie Jar and Sent Cookies

This section details the overall distribution of cookies collected during our measurement campaign across both popularity and random runs, as quantified in the numbers presented in Table 5.

Overall, the table classifies cookies under two main categories (databases): *Cookie Jar* and *Sent Cookies*. *Cookie Jar* refers to cookies that are set by websites in the stateful phase when their consent banners are accepted by the BannerClick tool [65]. *Sent Cookies* includes cookies extracted from the HTTP requests of websites before rejecting consent banners. Subsequently, the table's sub-columns categorize the cookies into different types, viz., 'Total,' 'First Party,' 'Third Party,' 'Tracking,' 'Intractable,' and 'Reset.' Specifically, 'Total' denotes the number of all cookies in the databases. 'Reset' refers to intractable cookies overwritten within the domain sending it, e.g., in Figure 1, `new.com` may also reset the intractable cookie (*id*=123).

For each run, the first row presents the cumulative count of cookies collected, reflecting the aggregated number of cookies set or sent across the domains. For instance, if a cookie is set twice by two different *setter websites*, it is counted as two. On the other hand, the second row shows the count of unique cookies grouped by their `name` and `host` attributes. Finally, the third row, labeled 'Avg,' depicts the mean number of cookies per domain.

Note that the table presents raw statistics without any adjustments. Across different runs, the number of accepted websites varies, influencing the number of cookies stored in the *Cookie Jar*. Runs with a greater number of accepted domains are likely to have a larger number of unique cookies in *Cookie Jar*, potentially resulting in a higher count of cookies sent per rejected domain. Consequently, a fair comparison cannot be made across the average numbers presented in the *Sent Cookies* columns of different runs. For a fair comparison, see § 5.

In comparing *Cookie Jar* of popularity runs, we observe that the average number of cookies is similar across all categories. In the case of random runs, the behavior across all categories appears relatively consistent. Additionally, in all runs, the counts of 'Unique' intractable cookies are identical between *Cookie Jar* and its corresponding *Sent Cookies*. This consistency is expected, as we classify a cookie in *Cookie Jar* as intractable only if it is subsequently sent from domains where consent is rejected.

Moreover, the average number of 'Reset' cookies is relatively low compared to the number of intractable cookies sent. This highlights the cohort nature of intractable cookies, as they tend to be sent via HTTP requests without being reset, making them difficult to track by simply observing the current state of cookies in the browser. It also implies that merely considering the deployment of tracking cookies overlooks a large portion of real-world tracking practices.

D Cookie Expiration and Renewal

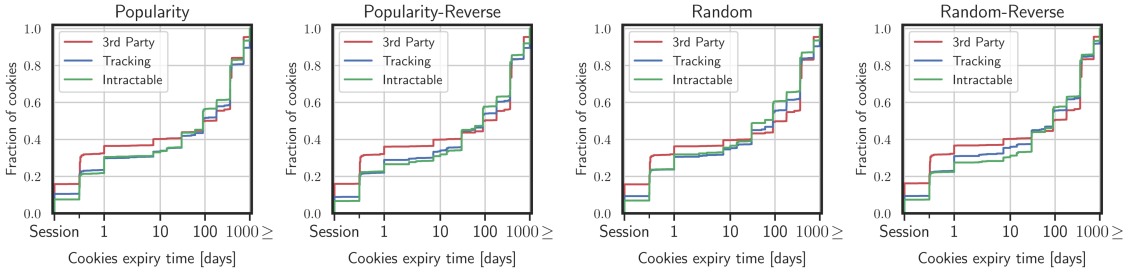
To further explore the expiration time and the number of times cookies are set, we analyze cookies categorized as third-party, tracking, and intractable for all four main runs. For the analysis, we use the unique cookies in a *Cookie Jar* (i.e., 'Unique' rows of Table 5).

Expiration: The ECDF graph displayed in Figure 13 illustrates the distribution of cookies according to their expiry time in days. The label 'Session' on the x-axis denotes cookies that are set to expire at the end of the browsing session. Across all categories and runs, we observe a consistent trend where nearly 60% of cookies have an expiry exceeding 10 days. Furthermore, we note that the most common expiry time among these cookies is 365 days.

Renewal: The ECDF graphs presented in Figure 14 illustrate the distribution of cookies based on the number of times they are set or reset across different websites. We observe that intractable cookies are more prone to being reset compared to tracking cookies, and even more so than third-party cookies, for all runs. As shown in Figure 13a, there is a subtle difference in the number of renewals of cookies across popularity runs. Specifically, the gap between graphs for different categories is larger in the *Popularity-Reverse* run compared to *Popularity* run. For instance, $\approx$ 30% of intractable cookies are set more than once in *Popularity* run, whereas this number is about 35% for the *Popularity-Reverse* run. Nevertheless, the trends show similar patterns across all runs.

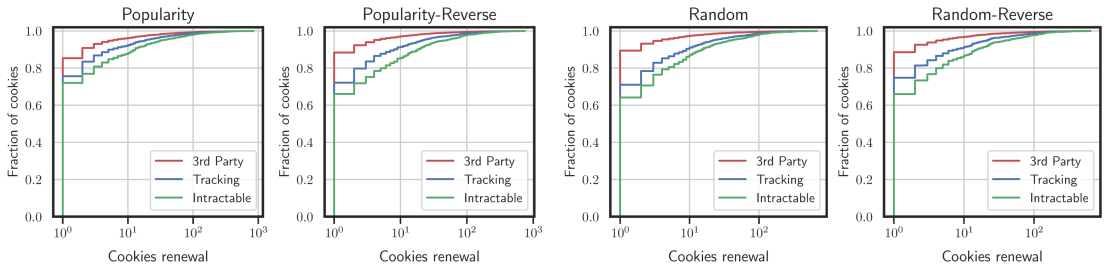
		Cookie Jar					Sent Cookies		
		Total	First Party	Third Party	Tracking	Intractable	Total	Intractable	Reset
Popularity	Aggregate	95,131	57,918	37,213	21,920	19,296	225,520	75,164	5,952
	Unique	67,083	57,782	9,301	3,583	2,131	31,239	2,131	530
	Avg	31.35	19.09	12.27	7.22	6.36	94.80	31.59	2.50
Popularity-Reverse	Aggregate	93,769	55,154	38,615	20,612	18,571	187,827	54,918	5,533
	Unique	67,195	55,115	12,080	3,108	1,769	28,935	1,769	530
	Avg	30.64	18.02	12.62	6.74	6.07	77.49	22.66	2.28
Random	Aggregate	84,266	49,315	34,951	17,610	16,041	214,779	59,605	6,591
	Unique	61,349	49,217	12,132	2,667	1,725	28,618	1,725	530
	Avg	28.73	16.81	11.92	6.00	5.47	83.31	23.12	2.56
Random-Reverse	Aggregate	83,718	51,180	32,538	17,338	15,455	201,862	68,040	6,820
	Unique	61,421	51,150	10,271	2,823	1,665	27,876	1,665	536
	Avg	28.07	17.16	10.91	5.81	5.18	79.38	26.76	2.68

Table 5: Cookie distribution across different measurements.



(a) Popularity and Popularity-Reverse runs **(b) Random and Random-Reverse runs**

Figure 13: Cookie Jar expiration time analysis.



(a) Popularity and Popularity-Reverse runs **(b) Random and Random-Reverse runs**

Figure 14: Cookie Jar renewal analysis.