Thou Shalt Not Reject: Analyzing Accept-Or-Pay Cookie Banners on the Web

Ali Rasaii
Max Planck Institute for Informatics
arasaii@mpi-inf.mpg.de

Devashish Gosain
BITS Pilani Goa
devashishg@goa.bits-pilani.ac.in

Oliver Gasser
Max Planck Institute for Informatics
oliver.gasser@mpi-inf.mpg.de

ABSTRACT

Privacy regulations have led to many websites showing cookie banners to their users. Usually, cookie banners present the user with the option to "accept" or "reject" cookies. Recently, a new form of paywall-like cookie banner has taken hold on the Web, giving users the option to either accept cookies (and consequently user tracking) or buy a paid subscription for a tracking-free website experience.

In this paper, we perform the first completely automated analysis of cookiewalls, i.e., cookie banners acting as a paywall. We find cookiewalls on 0.6% of all queried 45k websites. Moreover, cookiewalls are deployed to a large degree on European websites, e.g., for Germany we see cookiewalls on 8.5% of top 1k websites. Additionally, websites using cookiewalls send 6.4 times more third-party cookies and 42 times more tracking cookies to visitors, compared to regular cookie banner websites. We also uncover two large Subscription Management Platforms used on hundreds of websites, which provide website operators with easy-to-setup cookiewall solutions. Finally, we publish tools, data, and code to foster reproducibility and further studies.

CCS CONCEPTS

• **Networks** → **Network measurement**; **Network privacy and anonymity**.

KEYWORDS

Cookie banner, cookiewall, subscription management platform, Web measurement.

ACM Reference Format:
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), October 24–26, 2023, Montreal, QC, Canada.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3618257.3624846

1 INTRODUCTION

At first glance it may appear, that the vast majority of websites offer their content free of cost. However, many websites have an inherent cost for users by collecting their data and record their personal choices (e.g., in the form of cookies), which leads to targeted advertising. This entire user profiling and targeting nexus is sometimes referred to as "surveillance capitalism" [66]. To counter user tracking and safeguard user privacy, privacy laws such as the European Union's GDPR [23] have been enacted. They mandate that websites take explicit consent from users before storing or sharing their personal data. This led to an increase in cookie banners on websites. These banners notify users about personal data collection policies and provide interaction options like "accept" and "reject" to the users.

However, recently some websites have switched from showing regular cookie banners to using "cookiewalls". With cookiewalls users are given two options—either to provide consent for tracking or to buy a subscription to access a website's content without ads and tracking. GDPR clearly states that consent to the processing of personal data must be given freely and unconditionally [16]. Therefore, the legality of cookiewalls remains questionable and views of data protection authorities in European countries on the subject differ [43].

In this paper, we perform the first completely automated large-scale analysis of the cookiewall ecosystem to date. More specifically, the main contributions of this paper are:

- **Large-scale automated measurement study:** We perform a large-scale automated measurement study to detect cookiewalls from eight vantage points on 45k websites. We develop a tool to automatically detect cookiewalls with a precision of 98.2%, which we release as open-source [50] together with our analysis code and data [49] at bannerclick.github.io (see Section 3).
- **Characterization of cookiewall landscape:** We find cookiewalls on a total of 280 websites (0.6%), with some countries such as Germany seeing a 5 times higher prevalence with 2.9% of reachable websites (see Section 4.1). We analyze different cookiewall pricing schemes, finding that around 80% of websites charge 3 Euro per month or less (see Section 4.2). We investigate if buying a cookiewall subscription indeed protects from tracking and show that subscribers see no tracking cookies compared to an average of 16 tracking cookies seen by non-subscribers (see Section 4.3). We uncover that the majority of found cookiewall websites use Subscription Management Platforms to facilitate the deployment of cookiewalls (see Section 4.4). We highlight that common ad-blocking solutions are able to block 70% of cookiewalls using manually curated filter lists (see Section 4.5).
- **Discussion of cookiewall impact:** We discuss the impact of the advent of cookiewalls, reflect on dark patterns to compel users to accept tracking, and reason about the possible future of cookiewalls (see Section 5).

IMC '23, October 24–26, 2023, Montreal, QC, Canada

ACM ISBN 979-8-4007-0382-9/23/10.
https://doi.org/10.1145/3618257.3624846

2 BACKGROUND AND RELATED WORK

In recent years, various data protection laws such as the GDPR [23] in the EU or the CCPA [7] in California have been enacted to regulate the use of Web cookies and other tracking and profiling techniques. Although sharing similar goals, these laws are implemented in different forms. GDPR for example mandates that before any storage or exchange of personal information (e.g., cookies) the user needs to provide explicit informed consent (i.e., opt-in). CCPA on the other hand states that users must have the option to object to the sharing of their data (i.e., opt-out).

Many different scientific works have analyzed the efficacy of these laws [17, 55, 62], privacy policies [9, 18, 41, 46], third parties [35, 59], cookie banners [56, 64], and tracking in the Web [5, 21, 22, 24, 25, 30, 39, 40, 57].

To perform these studies in an automated way, different measurement tools—such as Selenium [58], OpenWPM [21, 65], FPdetective [2], Chameleon [6], and Common Crawl [11]—have been proposed. In 2021, Jha et al. [31] proposed Priv-Accept to accept cookie banners in an automated manner. In early 2023, Rasaii et al. [51] presented BannerClick, a tool which can automatically detect and interact (i.e., accept and reject) with cookie banners with an accuracy of 99% and 96%, respectively. Unfortunately, none of the currently available tools is able to automatically detect cookiewalls. Therefore, in this work we extend BannerClick to specifically detect these special types of cookie banners, i.e., cookiewalls.

Closest to our research are the works by Papadopoulos et al. [47] and by Morel et al. [43]. In the former paper, the authors investigate paywalls on websites and classify them into soft (limited number of articles can be read before the paywall is shown) and hard paywalls (a subscription is required to access content on a website). They identify 1.5k websites with some form of paywall-related JavaScript libraries on them. Contrary to our work, they do not look for cookiewalls on websites. In the latter paper, the authors manually annotate and classify cookiewalls on websites. They find 13 out of 2.8k websites (0.66%) showing a cookiewall to the user. In comparison to their work, we completely automate the task of detecting cookiewalls on websites, characterize the prevalent use of Subscription Management Platforms among cookiewalls for the first time, and conduct our study on a much larger set of target websites (more than 45k compared to 2.8k).

3 METHODOLOGY

In this section, we describe the vantage points and target domains for our measurements, detail our cookiewall detection approach, report on the accuracy of our technique, and discuss the limitations of our approach.

Vantage Points and Targets: We use AWS cloud instances at the following locations as our vantage points (VPs): Frankfurt (Germany), Stockholm (Sweden), Ashburn (US East), San Francisco (US West), Mumbai (India), São Paulo (Brazil), Cape Town (South Africa), and Sydney (Australia). We select these VPs as they include regions with different privacy regulations: GDPR in EU countries (Germany and Sweden), CCPA in California, and LGPD in Brazil. The remaining globally distributed VPs are in countries that have either no or less strict privacy regulations.

We use Google's Chrome User Experience Report (CrUX) [10] for target selection, as it was shown to be a more realistic toplist [54] compared to Alexa [3] or Tranco [38]. We take the union of the country-wise Google CrUX top 10k domains for each VP country, resulting in 45 222 unique domains reachable in all VPs.

Cookiewall Detection Approach: To measure the prevalence of cookiewalls, we use a heavily modified version of the tool BannerClick [51]. BannerClick is built on top of OpenWPM [21] and Selenium [58], and can automatically detect and interact with cookie banners on websites. We enhance BannerClick by adding support for HTML shadow DOMs [44] and implement a tailored technique to detect cookiewalls on websites.

In our tests, we find that multiple websites with cookiewalls use shadow DOM environments, which can not be directly modified or inspected by browsers or even Selenium [37] (e.g., it is not possible to look up elements inside shadow DOMs using XPath or CSS selectors). We work around this limitation by looking for possible elements within the main HTML DOM with the `shadow_root` property. Then we clone and append all child elements within a shadow DOM to the body element of the main document DOM. Thereafter we find the desired button in the cloned DOM and then run the interaction function on the corresponding element in the shadow DOM. This allows BannerClick to also detect and interact with banners within open and closed shadow DOMs [52].

Before detecting cookiewalls, we first run BannerClick to detect all types of cookie banners. We then leverage BeautifulSoup [53] to search for cookiewall-specific words and classify banners as cookiewalls. As cookiewalls provide a tracking-free website by paying a subscription fee, we assemble a corpus of cookiewall-specific words consisting of (1) words related to subscriptions (i.e., abo, abonnent, abbonamento, abonne, abonné, ad-free and subscribe) and (2) currency words and symbols[1]. For each currency word or symbol, we check for a possible payment-related combination, e.g., *$3.99*, *3.99$*, *3.99 $*, or *3.99 $*. If these combinations of currency words or cookiewall-related words appear in the text of a banner, we classify that banner as a cookiewall. In total, we find that out of 280 correctly detected cookiewall websites, 76 make use of a shadow DOM, 132 are embedded in iFrames, and 72 use the main HTML DOM to embed cookiewalls. In Appendix B we show example screenshots for cookiewalls and cookie banners. We release our modified version of BannerClick as open-source software [50].

Detection Accuracy: To measure the accuracy of our cookiewall detection approach, we randomly select 1000 domains from our target list and manually check their screenshots to find the possible existence of cookiewalls on the website. We find that we correctly detect all 6 present cookiewall websites. The remaining 994 websites indeed do not show a cookiewall. Therefore, for these 1000 random websites we have a precision and recall of 100%.

Furthermore, we manually check all 285 websites where we detected a cookiewall to gain confidence in our detection approach. We find that 280 websites have indeed a cookiewall, whereas 5 detections are classified as false positives. This results in a detection precision of 98.2%.

[1]We use the top 10 global currencies as well as the official currency of our measurement vantage points: EUR, USD, CHF, AUD, GBP, Rs, BRL, CNY, and ZAR.

VP	Cookiewalls	Toplist	ccTLD	Language
US East	197	0	0	9
US West	199	0	0	9
Brazil	196	0	0	0
Germany	280	259	233	252
Sweden	276	15	0	0
South Africa	199	0	0	0
India	192	0	0	10
Australia	190	5	0	10

Table 1: Number of detected cookiewalls depending on the country of the vantage point, country-specific toplist, TLD associated with that country, and the most commonly spoken language in that country.

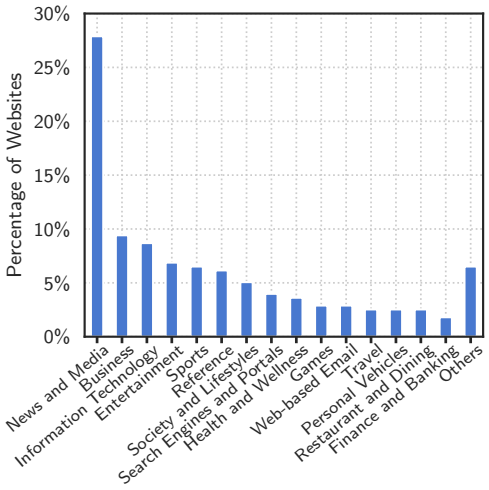

Figure 1: Categories of websites showing cookiewalls.

Limitations: Our study provides valuable insights into the prevalence and characteristics of cookiewalls. However, it is important to consider certain limitations when interpreting the results: First, we use an automated approach with a modified version of the BannerClick tool, achieving a 98.2% precision rate in detecting cookiewalls. However, false negatives are still possible, and manual verification may not guarantee complete accuracy for all websites. Second, some websites identify web crawlers as bots [36]. Thus when they detect a crawler, they may behave differently—e.g., altering the number of cookies or displaying cookiewalls differently from a regular user. Although OpenWPM has mechanisms to mitigate bot detection, it is impossible to completely circumvent bot detection. Hence, our study may not fully represent the actual website behavior experienced by regular users. Third, while our VPs are located in eight different geographical regions across six continents, more VPs in different countries can be added to the study. Thus, future studies can further increase the number of VPs across countries to obtain an even better understanding of cookiewalls. Finally, our study primarily examines the technical aspects and deployment of cookiewalls, not user perceptions or behaviors. Understanding user perspectives would require additional research, such as user surveys or studies.

4 MEASUREMENT RESULTS

In this section, we present results from our cookiewall measurements, including cookiewall prevalence across multiple characteristics, subscription pricing, third-party and tracking cookie analyses, a case-study of Subscription Management Platforms, and results from experiments to bypass cookiewalls.

4.1 Cookiewall Landscape

We use our modified version of BannerClick to run cookiewall measurements from eight vantage points targeting 45 222 websites. In Table 1 we show different characteristics of our measurements and the detected cookiewall websites. In total, we find cookiewalls on 280 unique websites, resulting in an overall cookiewall rate of 0.6%, a similar rate as found by previous work on a smaller set of target websites [43]. Our vantage points (VPs) in the EU (Germany and Sweden) see around 280 websites with cookiewalls compared to around 200 for non-EU VPs. This finding is consistent with the generally higher prevalence of cookie banners in the EU [51].

Next, we analyze different characteristics—i.e., country-specific toplists, top-level domains, and language—for *each vantage point* separately. We find that the Germany-specific CrUX toplist (see Section 3) contains by far the most detected cookiewall websites (259, 2.9% of reachable top 10k websites), followed by Sweden (15) and Australia (5). We also find cases where websites on a country-specific toplist show a cookiewall only when visited from a particular VP[2]. This shows that cookiewalls are affecting users differently based on the list of popular websites within their country.

To better understand websites showing cookiewalls to their visitors, we analyze the website top-level domain (TLD), the website's language, as well as the category the website can be attributed to. We find that again the vast majority of cookiewall websites are hosted on Germany's `.de` country-code TLD (ccTLD), followed by generic TLDs (14 on `.com`, 14 on `.net`, 4 on `.org`), and non-VP ccTLDs (6 on `.it`, 4 on `.at`, and 2 on `.fr`).

Next, we inspect the language of the cookiewall websites using CLD3 [26] to characterize the main target audience. Unsurprisingly, the largest part of these websites are in German[3], followed by English (US, Australia, India), Italian, and Swedish. To characterize the content of the website, we use FortiGuard's Web filter database [14] to assign each website to a category. As shown in Figure 1, more than one-fourth of all cookiewall websites are categorized as news and media, 9% fall into the business category, and 7% are IT-related websites. This highlights that cookiewalls—although they are most prominent on news websites—go beyond just news websites and are deployed on a large variety of different website categories.

[2]For example, the website pt.climate-data.org is on the Brazilian country-specific toplist, but only shows a cookiewall when visited from Germany or Sweden. This particular website is in fact operated by a German person, but provides specific subdomains for different languages, e.g., `pt.` for Portuguese.

[3]Note that this might also include websites targeted at readers outside Germany, e.g., Austria, Switzerland, or other German-speaking audiences.

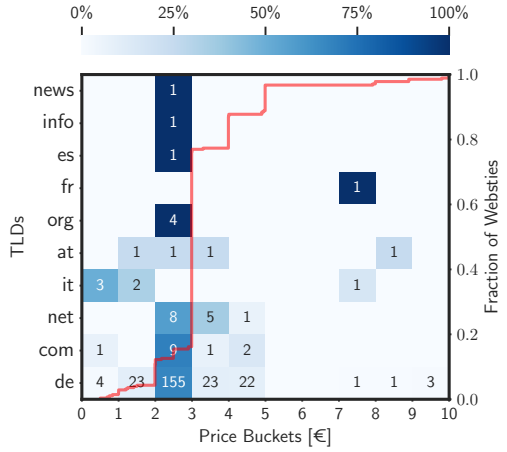

Figure 2: Distribution of monthly subscription price for cookiewall websites.

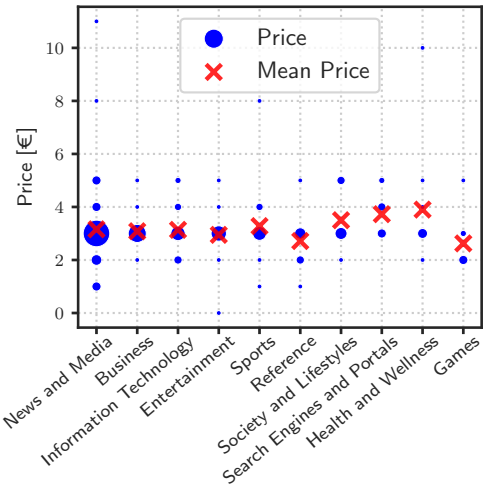

Figure 3: Correlation between the category of the websites and price of cookiewall website subscriptions.

Additionally, we find that cookiewalls are more prevalent on popular websites, i.e., 1.7% of country-wise top 1k domains show cookiewalls compared to 0.6% for top 10k domains[4]. Interestingly, if we just consider the top 1k reachable websites for Germany, we detect cookiewalls on more than 8.5% of websites, almost double the 4.7% in 2022 [43].

To summarize: *Cookiewalls are most prominent on websites which are popular among users from Germany, where we see them on 2.9% of top 10k websites and 8.5% of top 1k websites. Moreover, cookiewalls are visible on a wide variety of website categories, with news and media websites making up more than one fourth. In addition, more popular websites are more likely to show cookiewalls.*

4.2 Subscription Pricing

In this section, we analyze the price of cookiewall subscriptions of all detected 280 websites. We manually inspect each website to determine the price of a subscription. Then, we normalize the subscription price by month and convert it to Euro to make different websites comparable.

In Figure 2 we show the distribution of the monthly subscription price for cookiewall websites. The red line shows an ECDF for the prices of cookiewalls for all TLDs. We find that around 90% of cookiewall websites ask for 4 Euro (approx. 4.33 USD) or less per month, and by far the largest fraction of websites charges 3 Euro (3.25 USD), with the majority of these websites being attributed to a Subscription Management Platform in which subscribers just need to pay once to access all partnered websites (see Section 4.4). On the other end, a handful of websites ask for 9 Euro (9.74 USD) or more per month. The heatmap in Figure 2 shows the occurrence of each price bucket for each TLD separately. We find that TLDs of websites do not have a substantial impact on the prices, as most websites in different TLDs charge between 2 to 3 Euro per month, except for `.it` which are on average cheaper. Furthermore, we explore potential correlations between website categories and subscription prices. In Figure 3 the size of the blue data points represents the number of websites falling within each price range, with the red cross showing the mean price per category. We find no obvious relationship between subscription price and website category.

To summarize: *We find that 90% of cookiewall websites charge at most 4 Euro, with some outliers charging upwards of 9 Euro per month. Moreover, we find the prices to be generally similar for different TLDs and website categories.*

4.3 Third-party and Tracking Cookies

To assess the effect of cookiewalls on user privacy, we now compare cookies sent by websites with cookiewalls to websites with "regular" banners. Therefore, we run additional measurements targeting 280 cookiewall websites and 280 randomly selected websites with regular cookie banners with an accept button. To account for variations in advertisements and consequently sent cookies, we repeat each measurement five times per website and calculate the average number of cookies per website. We then compare the number of first-party, third-party, and tracking cookies after accepting cookiewalls and regular cookie banners. Similar to previous work [27, 51], we use the justdomains blocklist [32] to classify cookies as tracking cookies. If the cookie domain matches one of the domains in the justdomains list, we classify it as a tracking cookie. Note that there exist other techniques to track users that we do not consider in this research, e.g., browser fingerprinting [1], tracking using first-party cookies [8, 19, 45], and the use of invisible pixels and click IDs [4], as we specifically focus on studying the emergence of cookiewalls. Thus, in the future, a more nuanced analysis focusing on other tracking techniques can be conducted.

Figure 4 compares the average number of cookies set by websites with regular cookie banners and cookiewalls. In the figure, we see a similar number of first-party cookies among both website sets, with a median of 15 and 19 for regular cookie banner and cookiewall websites, respectively. In contrast, third-party cookies exhibit a

[4]Note that the Google CrUX toplist does not contain detailed rank information per website. It rather groups websites into rank buckets, e.g., top 1k or top 10k.

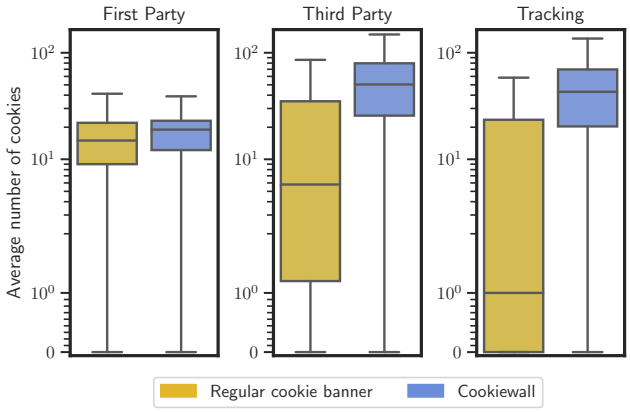

Figure 4: Average number of cookies comparing websites with regular cookie banners to cookiewall websites.

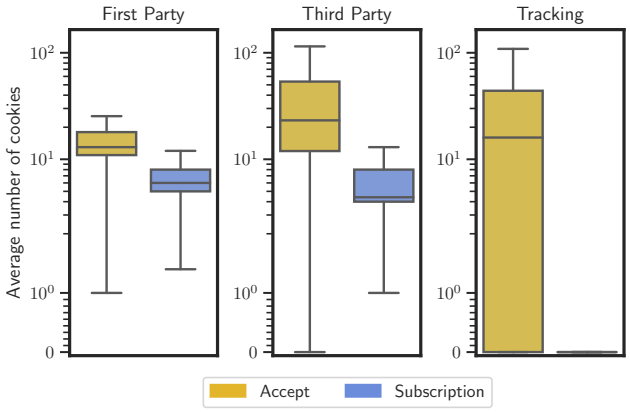

Figure 5: Average number of cookies set by websites with contentpass cookiewall after accepting or accessing with a subscription.

stark difference between both website sets. We find many more third-party cookies on cookiewall websites with a median of 50.4, compared to just 6.8 for cookie banner websites. An even more pronounced discrepancy can be seen for tracking cookies, with cookiewall websites sending on average 42 times more tracking cookies compared to cookie banner websites (median: 43 vs. 1). This seems to indicate, that websites with cookiewalls try to monetize their users more aggressively compared to other websites, either through subscription fees or excessive tracking and advertising.

To summarize: *Cookiewall websites send 6.4 times more third-party and 42 times more tracking cookies compared to "regular" cookie banner websites. This highlights the focus on monetization efforts of cookiewall websites.*

4.4 Subscription Management Platforms

Similar to Consent Management Platforms (CMPs) for regular cookie banners [29, 60], we find two different Subscription Management Platforms (SMPs) for cookiewalls: contentpass [13] and freechoice [61], which claim to host cookiewalls for 219 and 167 websites, respectively.[5] These two SMPs provide ad-free access to all partner websites for a monthly fee of 2.99 Euro. Note that only 76 contentpass and 62 freechoice partner websites are in our merged top 10k target list of previous measurements. We also find evidence of interoperability between CMPs and SMPs, with the CMP consentmanager providing integration support for the SMP contentpass [12].

In order to contrast the experience of subscribed users and users accepting tracking on SMP websites, we run an additional measurement for all 219 contentpass partner websites. Thus we create a contentpass account and buy a one-month subscription. We automate the login behavior on each of these websites and compare the sent cookies to clicking "accept". We again run five repetitions per website and average the number of cookies, in order to take website and advertisements variations into account.

Figure 5 shows the distribution of the average number of first-party, third-party, and tracking cookies across all 219 contentpass websites. We find a lower number of first-party (FP) and third-party (TP) cookies when accessing these websites with a subscription, with a median of 13 vs. 6 FP and 23.2 vs. 4.4 TP cookies for accept and subscription, respectively. The most apparent difference can be seen for tracking cookies, where we see no tracking cookies with a subscription compared to a median of 16 when accepting the cookiewall. Some websites send more than 100 tracking cookies when accessing these websites without a subscription. This underlines that cookiewall websites are in fact aggressively tracking users, likely to maximize their income from non-subscribing users via ads and to push them towards buying a subscription.

To summarize: *Subscription Management Platforms provide an easy way for website operators to monetize their users by offering them a subscription instead of being tracked and served with ads. While subscribed users see no tracking cookies, users accepting cookiewalls of the contentpass SMP see a median of 16 tracking cookies with some extreme cases sending more than one hundred tracking cookies.*

4.5 Bypassing Cookiewalls

This section delves into the feasibility, implications, and tools available for bypassing cookiewalls on websites. The forcible accept-or-pay scheme of cookiewalls might in the eyes of some users justify the act of bypassing it without being concerned about ethical considerations. One commonly employed method for bypassing cookiewalls is the use of ad-blocker browser extensions. Notable examples include "I don't care about cookies" [34] "Ninja Cookie" [15] and "uBlock Origin" [63]. In this section, we focus on investigating the effectiveness of uBlock Origin, one of the most popular ad-blocker extensions.

To evaluate its effectiveness, we conduct a measurement on our 280 detected cookiewall websites. We enable the uBlock Origin extension[6] and access each of the websites five times. We find that 196 (70%) websites no longer display cookiewalls across all iterations,

[5] We observe an increase in these numbers between May and September 2023 to 270 for contentpass and 184 for freechoice.

[6] We enable the by default disabled Annoyances filter lists to block cookiewalls.

while the remaining websites still exhibit the cookiewall prompt. Note that while browser extensions like uBlock Origin can effectively block resources with domains[7] listed in block lists (such as Easylist), they may not perfectly eliminate all types of cookiewalls. Some cookiewalls may be served locally or use lesser-known third-party domains, which could evade the blocking measures. Additionally, we manually inspect these 196 websites and find that all of them except two[8] work normally and do not show any ads.

To summarize: *Browser extensions like uBlock Origin can effectively block 70% of cookiewalls in our measurements.*

5 DISCUSSION

We now discuss the implications of our findings and present future research directions.

Paywalls vs. Cookiewalls: Existing research [47] reports on the rise of two types of Internet paywalls—hard and soft. With hard paywalls, users cannot access the website without first buying a subscription. With soft paywalls, users can freely view a certain number of articles before they need to buy a paid subscription. In this paper, we highlight the use of cookiewalls where users (1) have to pay to *not* opt-in to tracking or, (2) accept using a service with tracking, or (3) cannot access the website's content at all. From a monetary perspective, cookiewalls are similar to hard paywalls, but overall they adversely impact the clients' privacy. Due to this new "pay or get tracked" model, users may be conditioned to accept tracking cookies rather than paying for their privacy. This can result in privacy laws like GDPR being less effective. Moreover, in the future, websites may charge unreasonably high prices that could further compel users to accept tracking as their default choice. Although previous researchers [28, 60] also highlight the deployment of manipulative and non-compliant consent pop-ups by different CMPs, they do not consider cookiewalls. For instance, Toth et al. [60] report that CMPs like Quantcast provide configuration interfaces to set up cookie banners and restricted website access, i.e., limited (or no access) to website content before interacting with the banners.

Cookiewalls, Website Content, and Tracking Cookies: Websites that show cookiewalls may offer important content to their clients. We find that many websites showing cookiewalls are in top 1k of domains. Thus, users will either provide consent to tracking or pay to avoid it, as they would not want to cease access to the website content. Cookiewalls have the potential to create two classes of Web users: those that can afford to not being tracked, and those who need to pay for services with their data. In the future, user studies can be conducted to estimate the "monetary value of the content" on cookiewall websites.

Tracking cookies themselves are used to facilitate ad serving, thus bringing monetary value to the website. To see if there is a correlation between the number of tracking cookies a website sets for "accepting" users and the subscription price, we run an additional small experiment. As shown in Figure 6, we observe no meaningful linear correlation between the number of tracking cookies set by websites when accepting tracking and the subscription price.

[7]Example of patterns in the block lists which prevent further communication with CMPs to show the banners: `*cdn.opencmp.net/*`, `*consentmanager.net/*`, `*usercentrics.eu/*`.

[8]hausbau-forum.de detects uBlock and asks the user for deactivation. promipool.de is clickable but not scrollable.

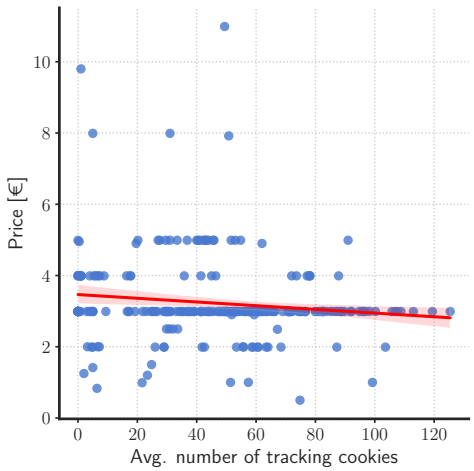

Figure 6: Correlation between the number of tracking cookies and price of cookiewall website subscriptions.

Circumventing Banners and Cookiewalls: Interaction with cookie banners could be seen as a nuisance to some users. Thus browsers such as Firefox are working on automatically clicking the reject button (if available) on banners [42]. This approach as well as our tool could lay the groundwork to also automatically interact with cookiewalls in the future.

Presently, we can use ad-block extensions and filter lists to evade most cookiewalls. There is, however, a risk that these extensions may also block necessary scripts, potentially disabling useful functionalities, or introducing security threats. Moreover, since ad-blockers run as a script on the client side, they can themselves be a source of privacy leaks.

Revoking Cookiewall Acceptance: We find that it is not trivial to switch from cookiewall acceptance to subscription. If a user has already consented to "accept" on some website's cookiewall, they must delete their cookies and local storage (specific to the website). After deletion, they would see the cookiewall on a subsequent visit and can change their choice. Since users will likely not be aware of these necessary additional steps, they might continue to be tracked even though they have subscribed e.g., on a different device.

6 CONCLUSION

In this paper, to the best of our knowledge we performed the first automated analysis of the cookiewall landscape to date. We developed a tool to automatically detect cookiewalls with a precision of 98.2%. Using this tool we crawled 45k websites and found cookiewalls on 280 of them. We investigated different cookiewall deployment characteristics and uncovered that they are especially deployed among popular websites in Germany (8.5%). Moreover, we compared cookiewalls to regular cookie banners and found websites to be sending 42 times more tracking cookies to cookiewall website visitors. Additionally, we uncovered two large Subscription Management Platforms which provide website operators with easily deployable cookiewall solutions. Finally, we publish our measurement tool to allow for future studies, as well as analysis code and data to foster reproducibility.

ACKNOWLEDGMENTS

We thank the anonymous reviewers as well as our shepherd Hamed Haddadi for their valuable feedback.

REFERENCES

[1] 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*. 674–689.

[2] Gunes Acar, Marc Juarez, Nick Nikiforakis, Claudia Diaz, Seda Gürses, Frank Piessens, and Bart Preneel. 2013. FPDetective: dusting the web for fingerprinters. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. 1129–1140.

[3] Amazon. 2023. Amazon Alexa. https://www.alexa.com/. (Accessed on 05/25/2023).

[4] Paschalis Bekos, Panagiotis Papadopoulos, Evangelos P Markatos, and Nicolas Kourtellis. 2023. The Hitchhiker's Guide to Facebook Web Tracking with Invisible Pixels and Click IDs. In *Proceedings of the ACM Web Conference 2023*. 2132–2143.

[5] Aaron Cahn, Scott Alfeld, Paul Barford, and Shanmugavelayutham Muthukrishnan. 2016. An empirical study of web cookies. In *Proceedings of the 25th international conference on world wide web*. 891–901.

[6] Chameleon Crawler contributors. 2015. Chameleon Crawler. https://github.com/ghostwords/chameleon.

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

[8] Quan Chen, Panagiotis Ilia, Michalis Polychronakis, and Alexandros Kapravelos. 2021. Cookie swap party: Abusing first-party cookies for web tracking. In *Proceedings of the Web Conference 2021*. 2117–2129.

[9] Rex Chen, Fei Fang, Thomas Norton, Aleecia M McDonald, and Norman Sadeh. 2021. Fighting the Fog: Evaluating the Clarity of Privacy Disclosures in the Age of CCPA. In *Workshop on Privacy in the Electronic Society (WPES)*.

[10] Chrome User Experience Report contributors. 2023. Chrome User Experience Report. https://developer.chrome.com/docs/crux/.

[11] Common Crawl. 2023. Common Crawl. https://commoncrawl.org/.

[12] consentmanager AB. 2023. Working with contentpass Integration. https://help.consentmanager.net/books/cmp/page/working-with-contentpass-integration.

[13] Content Pass GmbH. 2023. contentpass website. https://www.contentpass.net/.

[14] Fortiguard contributors. 2023. Web Filter Lookup | FortiGuard. https://www.fortiguard.com/webfilter. (Accessed on 05/25/2023).

[15] Ninja Cookie contributors. 2023. Ninja Cookie | Opt out of non-essential cookies and automatically remove cookie popups. https://ninja-cookie.com/. (Accessed on 05/26/2023).

[16] Cookie Banner Taskforce. 2023. Report of the work undertaken by the Cookie Banner Taskforce. https://edpb.europa.eu/system/files/2023-01/edpb_20230118_report_cookie_banner_taskforce_en.pdf.

[17] Adrian Dabrowski, Georg Merzdovnik, Johanna Ullrich, Gerald Sendera, and Edgar Weippl. 2019. Measuring cookies and web privacy in a post-gdpr world. In *International Conference on Passive and Active Network Measurement*. Springer, 258–270.

[18] 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 *Network and Distributed Systems Security (NDSS) Symposium*.

[19] Nurullah Demir, Daniel Theis, Tobias Urban, and Norbert Pohlmann. 2022. Towards Understanding First-Party Cookie Tracking in the Field. *arXiv preprint arXiv:2202.01498* (2022).

[20] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. 2013. ZMap: Fast Internet-wide Scanning and Its Security Applications. In *22nd USENIX Security Symposium (USENIX Security 13)*. 605–620.

[21] 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*. 1388–1401.

[22] Steven Englehardt, Dillon Reisman, Christian Eubank, Peter Zimmerman, Jonathan Mayer, Arvind Narayanan, and Edward W Felten. 2015. Cookies that give you away: The surveillance implications of web tracking. In *Proceedings of the 24th International Conference on World Wide Web*. 289–299.

[23] European Commission. 2023. The General Data Protection Regulation (GDPR) in EU. https://commission.europa.eu/law/law-topic/data-protection_en.

[24] Marjan Falahrastegar, Hamed Haddadi, Steve Uhlig, and Richard Mortier. 2014. The rise of panopticons: Examining region-specific third-party web tracking. In *International Workshop on Traffic Monitoring and Analysis*. Springer, 104–114.

[25] Roberto Gonzalez, Lili Jiang, Mohamed Ahmed, Miriam Marciel, Ruben Cuevas, Hassan Metwalley, and Saverio Niccolini. 2017. The cookie recipe: Untangling the use of cookies in the wild. In *2017 Network Traffic Measurement and Analysis Conference (TMA)*. IEEE, 1–9.

[26] Google. 2022. CLD3 on GitHub. https://github.com/google/cld3.

[27] Matthias Götze, Srdjan Matic, Costas Iordanou, Georgios Smaragdakis, and Nikolaos Laoutaris. 2022. Measuring Web Cookies in Governmental Websites. In *14th ACM Web Science Conference 2022*. 44–54.

[28] Colin M Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. 2021. Dark patterns and the legal requirements of consent banners: An interaction criticism perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.

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

[30] Costas Iordanou, Georgios Smaragdakis, Ingmar Poese, and Nikolaos Laoutaris. 2018. Tracing cross border web tracking. In *Proceedings of the Internet Measurement Conference 2018*. 329–342.

[31] Nikhil Jha, Martino Trevisan, Luca Vassio, and Marco Mellia. 2022. The Internet with Privacy Policies: Measuring The Web Upon Consent. *ACM Trans. Web* 16, 3, Article 15 (sep 2022), 24 pages. https://doi.org/10.1145/3555352

[32] justdomains. 2022. DOMAIN-ONLY Filter Lists. https://github.com/justdomains/blocklists.

[33] Erin Kenneally and David Dittrich. 2012. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. *Available at SSRN 2445102* (2012).

[34] Daniel Kladnik. 2023. I don't care about cookies. https://www.i-dont-care-about-cookies.eu/.

[35] Michael Kretschmer, Jan Pennekamp, and Klaus Wehrle. 2021. Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web. *ACM Trans. Web* 15, 4, Article 20 (July 2021), 42 pages. https://doi.org/10.1145/3466722

[36] Benjamin Krumnow, Hugo Jonker, and Stefan Karsch. 2022. How gullible are web measurement tools? A case study analysing and strengthening OpenWPM's reliability. In *Proc. 18th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '22)*. ACM, New York, NY, USA, 16. https://doi.org/10.1145/3555050.3569131

[37] Lavanya. 2023. How can we find the Xpath for Shadow Element. https://www.numpyninja.com/post/how-can-we-find-the-xpath-for-shadow-element. (Accessed on 05/26/2023).

[38] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium*. Internet Society, 1–15.

[39] Adam Lerner, Anna Kornfeld Simpson, Tadayoshi Kohno, and Franziska Roesner. 2016. Internet Jones and the Raiders of the Lost Trackers: An Archaeological Study of Web Tracking from 1996 to 2016. In *25th USENIX Security Symposium (USENIX Security 16)*. USENIX Association, Austin, TX. https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/lerner

[40] Tai-Ching Li, Huy Hang, Michalis Faloutsos, and Petros Efstathopoulos. 2015. Trackadvisor: Taking back browsing privacy from third-party trackers. In *International Conference on Passive and Active Network Measurement*. Springer, 277–289.

[41] Thomas Linden, Rishabh Khandelwal, Hamza Harkous, and Kassem Fawaz. 2020. The Privacy Policy Landscape After the GDPR. *Proceedings on Privacy Enhancing Technologies* 2020, 1 (2020).

[42] Martin Brinkmann. 2023. Firefox may soon reject Cookie prompts automatically. https://www.ghacks.net/2023/04/17/firefox-may-interact-with-cookie-prompts-automatically-soon/.

[43] Victor Morel, Cristiana Santos, Yvonne Lintao, and Soheil Human. 2022. Your Consent Is Worth 75 Euros A Year—Measurement and Lawfulness of Cookie Paywalls. In *Proceedings of the 21st Workshop on Privacy in the Electronic Society*. 213–218.

[44] Mozilla. 2023. MDN: Using shadow DOM. https://developer.mozilla.org/en-US/docs/Web/Web_Components/Using_shadow_DOM.

[45] Shaoor Munir, Sandra Siby, Umar Iqbal, Steven Englehardt, Zubair Shafiq, and Carmela Troncoso. 2023. CookieGraph: Understanding and Detecting First-Party Tracking Cookies. In *ACM Conference on Computer and Communications Security (CCS) 2023*.

[46] Sean O'Connor, Ryan Nurwono, Aden Siebel, and Eleanor Birrell. 2021. (Un) clear and (In) conspicuous: The right to opt-out of sale under CCPA. In *Workshop on Privacy in the Electronic Society (WPES)*.

[47] Panagiotis Papadopoulos, Peter Snyder, Dimitrios Athanasakis, and Benjamin Livshits. 2020. Keeping out the masses: Understanding the popularity and implications of internet paywalls. In *Proceedings of The Web Conference 2020*. 1433–1444.

[48] Craig Partridge and Mark Allman. 2016. Ethical considerations in network measurement papers. *Commun. ACM* 59, 10 (2016), 58–64.

[49] Ali Rasaii. 2023. Analysis scripts and raw data for Cookiewall measurements . https://doi.org/10.17617/3.TREBZR.

[50] Ali Rasaii. 2023. BannerClick on GitHub. https://github.com/bannerclick/bannerclick.

[51] Ali Rasaii, Shivani Singh, Devashish Gosain, and Oliver Gasser. 2023. Exploring the Cookieverse: A Multi-Perspective Analysis of Web Cookies. In *Proceedings of*

the 2023 Passive and Active Measurement Conference. https://doi.org/10.1007/978-3-031-28486-1_26
[52] Leon Revill. 2017. Open vs. Closed Shadow DOM. https://blog.revillweb.com/open-vs-closed-shadow-dom-9f3d7427d1af.
[53] Leonard Richardson. 2007. Beautiful soup documentation. *April* (2007).
[54] Kimberly Ruth, Deepak Kumar, Brandon Wang, Luke Valenta, and Zakir Durumeric. 2022. Toppling top lists: evaluating the accuracy of popular website lists. In *Proceedings of the 22nd ACM Internet Measurement Conference, IMC 2022, Nice, France, October 25-27, 2022*, Chadi Barakat, Cristel Pelsser, Theophilus A. Benson, and David R. Choffnes (Eds.). ACM, 374–387. https://doi.org/10.1145/3517745.3561444
[55] Iskander Sanchez-Rola, Matteo Dell'Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos. 2019. Can i opt out yet? gdpr and the global illusion of cookie control. In *Proceedings of the 2019 ACM Asia conference on computer and communications security*. 340–351.
[56] Cristiana Santos, Arianna Rossi, Lorena Sanchez Chamorro, Kerstin Bongard-Blanchy, and Ruba Abu-Salma. 2021. Cookie Banners, What's the Purpose? Analyzing Cookie Banner Text Through a Legal Lens. In *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society* (Virtual Event, Republic of Korea) *(WPES '21)*. Association for Computing Machinery, New York, NY, USA, 187–194. https://doi.org/10.1145/3463676.3485611
[57] Sebastian Schelter and Jérôme Kunegis. 2016. Tracking the trackers: A large-scale analysis of embedded web trackers. In *Tenth International AAAI Conference on Web and Social Media*.
[58] Selenium. 2023. Browser automation using Selenium. https://www.selenium.dev/.
[59] Jannick Sørensen and Sokol Kosta. 2019. Before and after gdpr: The changes in third party presence at public and private european websites. In *The World Wide Web Conference*. 1590–1600.
[60] Michael Toth, Nataliia Bielova, and Vincent Roca. 2022. On dark patterns and manipulation of website publishers by CMPs. *Proceedings on Privacy Enhancing Technologies* 3 (2022), 478–497.
[61] Traffective GmbH. 2023. freechoice website. https://freechoice.club/.
[62] Martino Trevisan, Stefano Traverso, Eleonora Bassi, and Marco Mellia. 2019. 4 Years of EU Cookie Law: Results and Lessons Learned. *Proc. Priv. Enhancing Technol.* 2019, 2 (2019), 126–145.
[63] uBlock Origin contributors. 2023. uBlock Origin - Free, open-source ad content blocker. https://ublockorigin.com/. (Accessed on 05/26/2023).
[64] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (un) informed consent: Studying gdpr consent notices in the field. In *Proceedings of the 2019 acm sigsac conference on computer and communications security*. 973–990.
[65] WebTAP at Princeton University. 2023. Studies using OpenWPM. https://webtap.princeton.edu/software/.
[66] Shoshana Zuboff. 2019. Surveillance capitalism and the challenge of collective action. In *New labor forum*, Vol. 28. SAGE Publications Sage CA: Los Angeles, CA, 10–29.

A ETHICS

We incorporate proposals by Partridge and Allman [48] and Kenneally and Dittrich [33] and follow best measurement practices [20] when running our measurements. We use dedicated measurement machines, set up informative rDNS names, host a website with information about our measurements, and offer the possibility to be blocklisted from the measurements. We run OpenWPM in a similar way as any regular user when visiting websites with a normal Web browser. During our measurement period, we did not receive any complaints.

B SCREENSHOTS

Figure 7 shows a screenshot of an example cookiewall on a website, whereas Figure 8 shows a screenshot of a regular cookie banner on a website. Note the presence of a subscription button instead of a "reject" or "options"/"manage my cookies" button in the cookiewall.

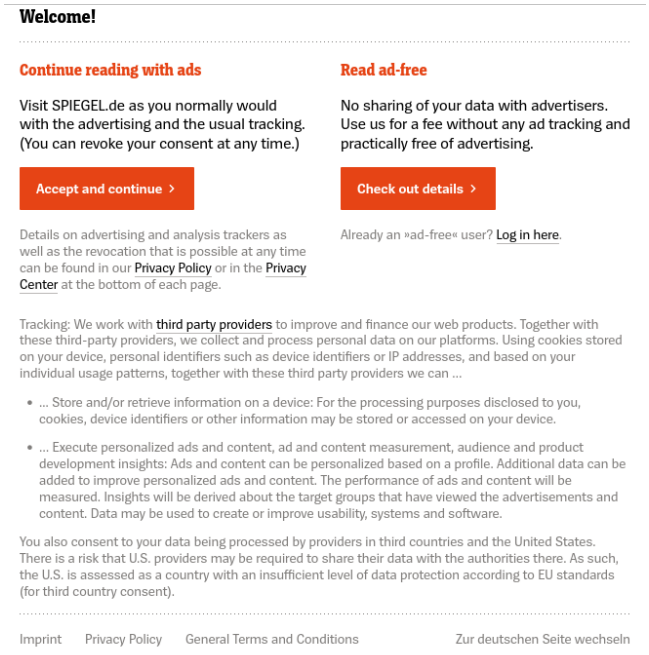

Figure 7: Example of a cookiewall shown on spiegel.de.

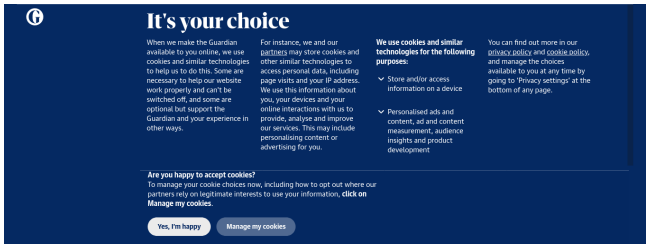

Figure 8: Example of a regular cookie banner shown on guardian.co.uk.