Measuring Cookies and Web Privacy in a Post-GDPR World

Adrian Dabrowski[1][0000−0002−0340−6204],
Georg Merzdovnik[1][0000−0002−9955−7284],
Johanna Ullrich[2][0000−0003−0297−9614],
Gerald Sendera[1], and Edgar Weippl[2][0000−0003−0665−6126]

[1] SBA Research, Vienna, Austria
{adabrowski,gmerzdovnik,gsendra}@sba-research.org
http://www.sba-research.org
[2] Christian Doppler Laboratory for Security and Quality Improvement in the Production System Lifecycle, Institute of Information Systems Engineering, TU Wien, Austria
{johanna.ullrich,edgar.weippl}@tuwien.ac.at

Abstract. In response, the European Union has adopted the General Data Protection Regulation (GDPR), a legislative framework for data protection empowering individuals to control their data. Since its adoption on May 25th, 2018, its real-world implications are still not fully understood. An often mentioned aspect is Internet browser cookies, used for authentication and session management but also for user tracking and advertisement targeting.
In this paper, we assess the impact of the GDPR on browser cookies in the wild in a threefold way. First, we investigate whether there are differences in cookie setting when accessing Internet services from different jurisdictions. Therefore, we collected cookies from the Alexa Top 100,000 websites and compared their cookie behavior from different vantage points. Second, we assess whether cookie setting behavior has changed over time by comparing today's results with a data set from 2016. Finally, we discuss challenges caused by these new cookie setting policies for Internet measurement studies and propose ways to overcome them.

Keywords: GDPR · Cookies · Privacy.

1 Introduction

Privacy means freedom from (unauthorized) surveillance and is considered a human right according to the United Nations. Nowadays, individual privacy is increasingly eroding by our fully digitalized world: Enterprises of the digital economy such as social networks or online advertisers but also nation-state actors collect vast amounts of data via the Internet [8].

However, privacy legislation significantly differs among countries failing to address the international aspects of the Internet. With the General Data Protection Regulation (GDPR), the European Union (EU) has taken the effort of

harmonization and adopted bold rules for personal data. Its overall goal is to empower individuals to control their personal data and is largely considered one of the strictest legislative framework for data protection worldwide [26]. As a novelty, the GDPR does not only apply to data collection and processing based in the EU, but also to data of European residents that is collected abroad. In theory, this implies changes to all Internet services offered to EU customers regardless of their origin.

As one of its aims, GDPR seeks to prevent the (unconsenting) creation of user profiles; in consequence, most common usages of browser cookies are affected.

As a reaction, Internet services appear to follow one of the strategies below for compliance with the GDPR: (1) A service refrains from using persistent cookies at all. (2) The service asks for explicit user consent and only then sets the cookies, leaving the site usable without consent. In practice, there is frequently a banner spanning over a service's pages asking for consent. (3) Alternatively, EU users are banned from using the service. For example, Los Angeles Times [16] remained inaccessible from Europe for some time, and even *GDPR shields as a service* preventing visits from Europe for a monthly fee are available [13].

In this paper, we assess the impact of the recent GDPR enactment on cookie setting behavior at large scale. In particular, our research is threefold:

- We investigate whether there are differences in cookie setting when accessing the Alexa Top 100,000 websites from different jurisdictions by collecting cookies in an Internet measurement study. We compare persistent cookie usage upon requests originating from the European Union with such from the United States.
- Further, we assess whether cookie setting behavior has changed with the implementation of the GDPR in May 2018. Therefore, we compare our 2018 results with a data set collected in 2016.
- Finally, we infer challenges for Internet measurement studies imposed by GDPR's implementation and discuss means to overcome them.

The remainder of the paper is organized as follows: Section 2 provides background on browser cookies and their subsumption under the GDPR framework. Section 3 describes our measurement methodology and data sets. Our results on differences in cookie setting behavior with respect to jurisdiction are presented in Section 4, on changes over time in Section 5. Section 6 discusses the impact of cookie setting policies on Internet-wide measurement studies, and Section 8 concludes.

2 Background and Related Work

This section provides background on HTTP cookies, the General Data Protection Regulation (GDPR) and finally presents related work.

2.1 HTTP Cookies

HTTP cookies are a state management mechanism [2], enabling stateful behavior for the per se stateless HTTP protocol [12]. Cookies are data pieces – to be

precise: name-value pairs and metadata – and are set (1) in a server's HTTP response using a HTTP `Set-Cookie` header or alternatively (2) by Javascript running at the client. Either way, the cookie is stored in the client's browser and sent back to the server in subsequent requests; whether the cookie is included into a request is decided upon its metadata, e.g., its expiry date, domain and path. The server is able to adapt its behavior in dependence of the cookie information, or return a modified cookie to the client. Cookies have manifold usages, including user authentication, user tracking or targeted advertising. The latter two are typically permanent (or persistent) cookies; they are issued automatically at the first visit to a new site, stored non-volatile, remain valid for long periods of time – as a consequence of the chosen expiry date – and are used to link subsequent visits of the same user. In contrast, cookies for login purposes are only set during authentication. Working as a temporary session identifier, they are only kept in volatile memory, expire within hours and lost when the browser window is closed.

2.2 General Data Protection Regulation

The General Data Protection Regulation (GDPR) came into force on May 25th, 2018. Before that, cookies have been addressed in 2002 by the European Union's ePrivacy directive (Directive 2002/58/EC, Directive on privacy and electronic communications), also known as *Cookie directive*, and its adaption in 2009 (2009/136/EC). Publications of the European Commission address the correct use of cookies in regard of the European legal framework [10]. In case of cookies acting as a means to collect data for behavioral analytics or to facilitate user tracking, the legislation of the GDPR applies in addition.

According to the GDPR, a data controller will need a legal basis to process personal identifiable information (PII) at all. In general, this is either consent from the user or one of the exceptions in Article 6 of the GDPR.

1. *the data subject has given consent to the processing of his or her personal data for one or more specific purposes.* Almost any processing activity can be justified by informed consent, although the requirements for the underlying information are relatively high. For consent to be valid, it must be informed, specific, freely given and must constitute a real indication of the individual's wishes[3]. In any case, consent is only valid in the form of an active act of consent, or a so-called *opt-in*, the mere *opt-out* or tolerance would no longer be legal. An example in which consent would be required are behavioral analyses, e.g., by Google Analytics. Cookies are only exempted from this requirement, if they are used for the sole purpose of carrying out the transmission of a communication, and are strictly necessary in order for the provider of an information society service explicitly required by the user to provide that service.
2. *processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject*

[3] *Guidelines on Consent under Regulation 2016/679.* Adopted on 28 November 2017

prior to entering into a contract. This applies to cases in which the cookie is necessary to provide the requested webpage or service, such as cookies for user authentication, or webshop carts.

3. *processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data.* IT security, i.e. the protection from attacks or malware are considered a possible legitimate interest on the part of the person responsible.

Also in the case of transfer to a third country according to Article 44, which is often the case when using analysis, content delivery or social media (e.g. Google, USA), certain regulations must be observed in order to remain within a legal framework [11], e.g., in the case of the U.S. the Privacy Shield framework [9].

2.3 Related Work

Previous measurement studies collected cookies for various purposes; measurements are either active using crawling or probing a data set, e.g., [3, 5, 7, 8] or passive relying on (already available or just captured) traffic logs, e.g., [14, 25]: Hannak et al. [15] investigate price discrimination and personalization that is based on cookies. Englehardt et al. [8] studied the potential of passive mass surveillance, e.g., by intelligence services, exploiting third-party HTTP tracking cookies. Englehardt and Narayanan [7] developed a measurement platform for web privacy measurements. The platform allowed to assess cookie-based, stateful as well as fingerprinting-based, stateless tracking, and found previously unknown techniques for tracking. Merzdovnik et al. [21] provided a similar large scale study on tracking and tracker blocking, which was conducted on the Alexa Top 200.000 pages. Cahn et al. [3] crawled the Alexa Top 100,000 web sites for cookies; their assessment led to the development of a mathematical model to quantify user information leakage. Sivakorn et al. [25] assessed the extent of information that is revealed to adversaries hijacking HTTP cookies in cases of simultaneous deployment of HTTP and HTTPS. Based on a data set of cookies collected in the wild, Gonzalez et al. [14] developed a methodology extracting information from proprietary data formats that are frequently used in cookies. Lerner et al. [19] investigated the history of cookie-based web tracking from 1996 to 2016; their work is similar to ours insofar as historic changes in cookie utilization are observed. We progress this line of research by specifically assessing the practical impact of the GDPR adoption on cookie usage.

Degeling et al. [6] and Linden et al. [20] made studies on the related topic of how GDPR changes web sites' privacy policies. Kulyk et al. [18] observed user reaction to the former cookie information banners.

Beyond, there are works on effects of GDPR adoption on Internet measurement: Plonka and Berger [22] revised collection of IPv6 addresses – the latter considered as personal identifiable information (PII) under the GDPR – and proposed a method to anonymize data sets of IPv6 addresses. Trammel and

Kühlewind [27] showed that even simple round-trip measurements reveal sensitive information, i.e., geographical location, and might thus become subject to the GDPR. While these works consider the impact of GDPR on Internet measurement, they do not assess the impact of GDPR adaption in the wild.

3 Methodology and Data Sets

For our study we employed three measurements based on the Alexa Top 100,000 ranking (a subset of Alexa Top 1 Million [1]). While the ranking methodology is poorly documented [23, 24], it remains a de-facto industry standard. In our opinion, it still depicts the World Wide Web better, than the protocol-agnostic Umbrella [4] list.

Two data sets were recorded in 2018 from different locations starting one week after GDPR's adoption. The third one had been recorded in 2016 but remapped to fit the 2018 ranking based on domain name. Thus, all data are shown and compared in 2018 ranking.

3.1 2018 Measurement

We crawled the Alexa Top 100,000 websites using the official Google Chrome v.66 browser in headless and network-deterministic mode. Every website was visited using HTTP and HTTPS, and each visit consisted of three retrievals. In the first retrieval, the main landing page is gained. For example, `microsoft.com` redirects to a a specific regional page such as `https://www.microsoft.com/nl-nl/` and `de.wikipedia.org` redirects to `https://de.wikipedia.org/wiki/Wikipedia:Hauptseite`. If the HTTP visit redirected to HTTPS or vice-versa, the other visit was skipped. The last two retrievals shared the user profile, but the browser process was closed in between. Thus removing all session cookies and leaving only persistent cookies. Finally, the internal *netlog* of the third retrieval was searched for transmitted cookies to the landing page. Unresponsive websites (e.g., due to down times or blocked IPs) and unresolvable websites were marked as erroneous.

This approach provides the following benefits: (1) The final netlog is not cluttered with redirect chains, e.g., HTTP redirects to HTTPS followed by further redirects to regional language pages. (2) In consequence, the first request in the netlog is the page actually shown to the user. (3) This way, subsequent requests to third party resources (e.g., advertisement banners) can be clearly separated. (4) Since persistent cookies can also be set dynamically among others by Javascript or in iframes, we have to observe the actual network traffic of an request to gain all cookies. It is not enough to look for `Set-Cookie` headers in server responses as the dynamic cookies are missed this way.

For the two measurements, we visited the Alexa Top 100,000 websites[4] from two locations. The first is based in a member state of the EU, the second in a

[4] We used Alexa Top 1 Million file [1] dating, May 24th, 2018, for all 2018 measurements after the introduction of GDPR.

US-based Amazon data center (east coast). Unreachable or unsuccessfully loaded websites were retried a few days later before being eventually marked as erroneous.

3.2 2016 Measurement

The 2016 measurement was conducted for a different study and thus used a slightly different approach. A scripted Firefox headless browser visited the Alexa Top 200,000 pages[5] from a US-based Amazon data center (west coast). These 200,000 measurements where later mapped to the 2018 ranking for comparison, matching 62,679 sites. A visit consisted of following initial redirects and then visiting three random sub pages of the site. Both, the HTTP and HTTPS traffic was recorded using a man-in-the-middle proxy with a custom-supplied certificate which was trusted by the browser.

This method is able to distinguish between session and persistent cookies that are set in HTTP headers using the expiry date from the `Set-Cookie` HTTP response header. However, it cannot make such a distinction on cookies that have been dynamically set (Javascript). The latter are only observed in subsequent requests but do not carry any indicator about their lifespan. In consequence, we can only provide an upper and lower bound for persistent cookies within this work representing the uncertainty span.

3.3 Unavailable Websites

In 2018, 2.5% of all websites were unavailable from the US but available from the EU; 0.7% were unavailable from the EU but available from the US and 1.9% were unavailable from both locations. Possible reasons include:

- Content delivery network (CDN) domains without an actual website. Examples include twimg.com (the Twitter CDN for images), cdninstagram.com, cloudfront.net (the CDN domain of cloudfront.com), ytimg.com, ebaydesc.com (used for user-supplied content to mitigate SOP-based problems such as XSS).
- Geographic blocking. For example, 92 Russian, 45 Iranian, and 58 Brasilian websites were exclusively blocked for US visitors, compared to 26, 5, and 13 respectively for EU visitors (based on ccTLD).
- Network and server failures during the measurement period.

3.4 Data Set Usage

For the 2018 measurements, we concluded that a website uses persistent cookies whenever at least one cookie was detected during the first HTTP(S) request in the third retrieval. We did not specifically look at third-party resources, such as advertisement networks. For the 2016 data set, we counted the `Set-Cookie`

[5] Alexa file from May 24th, 2016

header in case it included an expiry date or if a new cookie not previously set by HTTP headers was observed in the traffic. For comparison of the 2018 data sets, we used websites reachable from both measurement locations only. For temporal analysis, we matched the 2018 and 2016 data sets by domain name and included only domains that have been prevalent in both data sets since the Alexa ranking changed considerably over time.

3.5 Ethical Considerations

In our measurements, no personal data of individuals was collected or processed, i.e., the GDPR does apply to our measurements. We pretended to be a regular user, and investigated how websites react in case of such legitimate use. However, personal data of real people was not involved at any stage of this study. Beyond, we did not inflict any negative impact on websites ranked in the Alexa ranking [1] as our line of action, i.e., accessing websites and storing the cookies in our browser, is considered regular behavior on the Internet.

4 Geographical Differences

In this section, we compare the results from our 2018 measurements from US and EU vantage points (Figure 1). In total, 94,836 sites could be retrieved from both

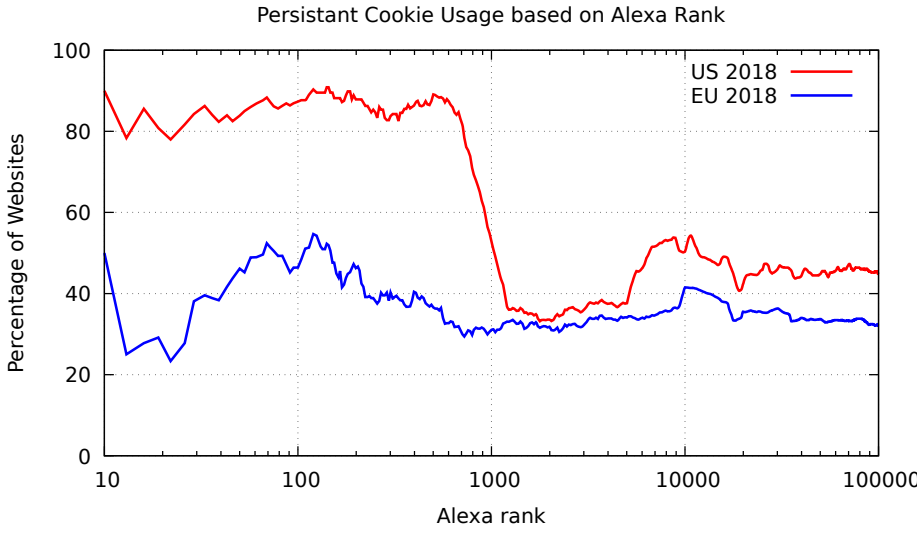

Fig. 1. Comparison of 2018 cookie usage.
How to read: Similar to a density plot. Example: Around rank position 100, 87% of sites install a persistent cookie when visited from the US, but only 46% sites do the same for European visitors without prior consent.

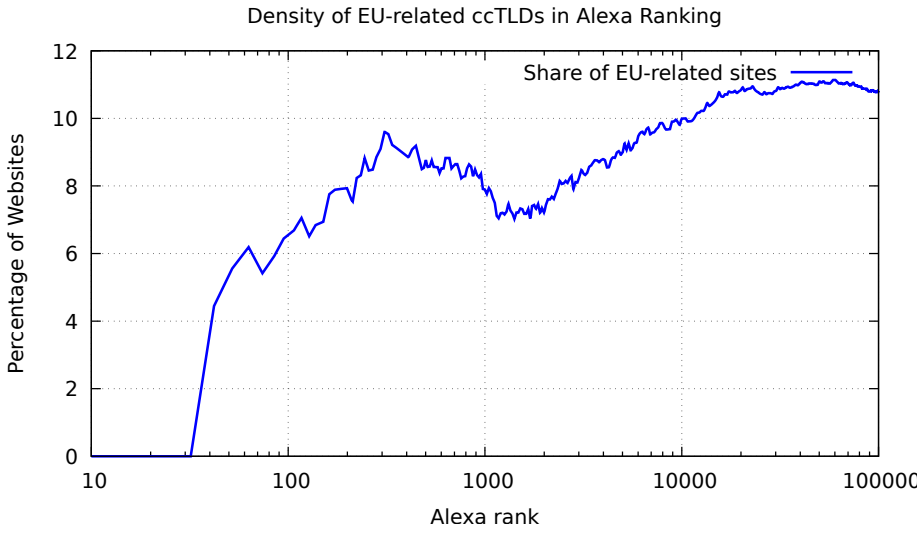

Fig. 2. Share EU-related top-level domains in the Alexa top 100,000 sites as of May 2018 based on ccTLD (Section 5.1)

test locations. Thereof, 50,663 sites (53.4%) did not install a persistent cookie on first visit, neither when requested from the EU nor the US. 31,362 (33.0%) installed a cookie in every case. However, 11,773 (12.4% of reachable and 26.6% of cookie-using) sites issued cookies for US-based visitors but not for EU-based ones. In comparison, only 1,038 (1% reachable, 2.3% cookie-using) of the sites set cookies for European, but not for US customers.

Interestingly, the discrepancy sharply increases for the top 1,000 websites: 49.3% of cookie-using websites choose to evade GDPR by using some form of geographic discrimination.

Figure 1 depicts the percentage of websites using persistent cookies, and shows the clear tendency of less persistent cookie usage for EU users. We have chosen a logarithmic scale for the Alexa rank as it better fits the long-tail characteristic of the rank.

We attribute the clear tendency to more geographically-based cookie differentiation on the higher ranks to the dominance of commercial and international websites. The Alexa list (and probably the WWW as a whole) is skewed towards non-EU sites on the top ranks (Figure 2). As EU-based operators have to apply GDPR rules regardless of origin, those rules have larger effect on lower ranks. Additionally, lower ranked websites tend not to have made the same investment into a real-time geographical differentiation of their visitors, as higher ranked websites. The latter's business model might depend more on tracking, advertisement and user analytics.

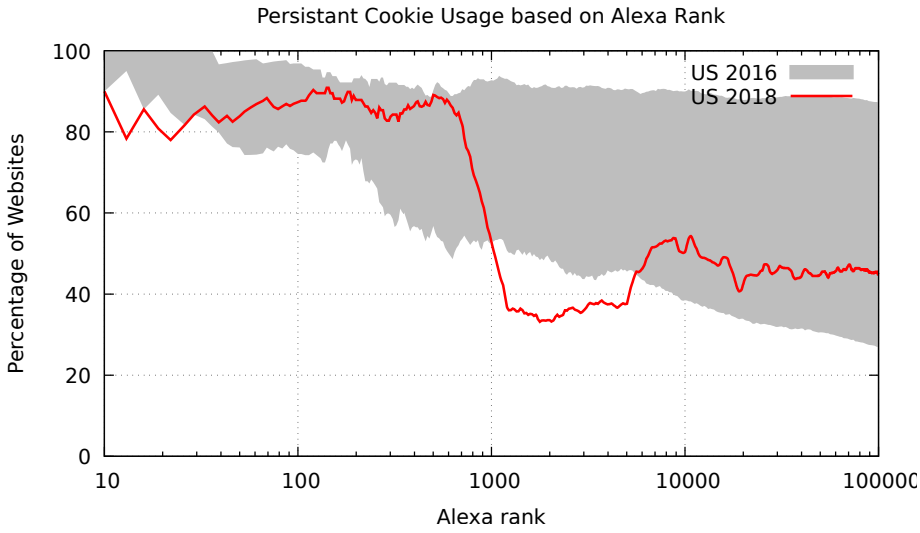

Fig. 3. Temporal comparison of cookie usage. Data for 2016 is given as an upper and lower limit (see Section 3.2) and mapped to Alexa ranking of May 2018

5 Temporal Changes

For temporal comparison, we normalized the 2016 ranking to fit the 2018 ranking. Thus, giving us a comparison about the development of cookie usage on a per-site basis. As for the above described uncertainty of the 2016 data, we can only provide ranges. The results are depicted in Figure 3.

In the top 1,000 ranks, 84% of sites could be matched from the whole 2016 data set. Of those, between 10.3% and 24.3% dropped non-consensual persistent cookie usage between 2016 and 2018.

In comparison, for the top 100,000 sites, only 55% could be matched from 2016 to 2018. Between 30.88% and 46.7% decided to refrain from non-consensual persistent cookie use all together.

Overall, US customers seem to also have profited from EU restrictions on Cookie usage.

As mentioned in Section 4, the drop of cookie usage above rank 1,000 for non EU-consumers could be due to websites applying EU-sane settings without a geographical considerations.

5.1 EU websites

Out of the Alexa Top 100,000 sites, we estimate are 10,584 are EU-related based on the country-code top-level domain (ccTLD)[6]. 7,668 have been valid

[6] eu,at,be,bg,cz,cy,de,dk,ee,es,fi,fr,gg,uk,gb,gi,gr,hr,hu,ie,im,it,je,lt,lu,lv,mt, nl,pl,pt,ro,se,si,sk

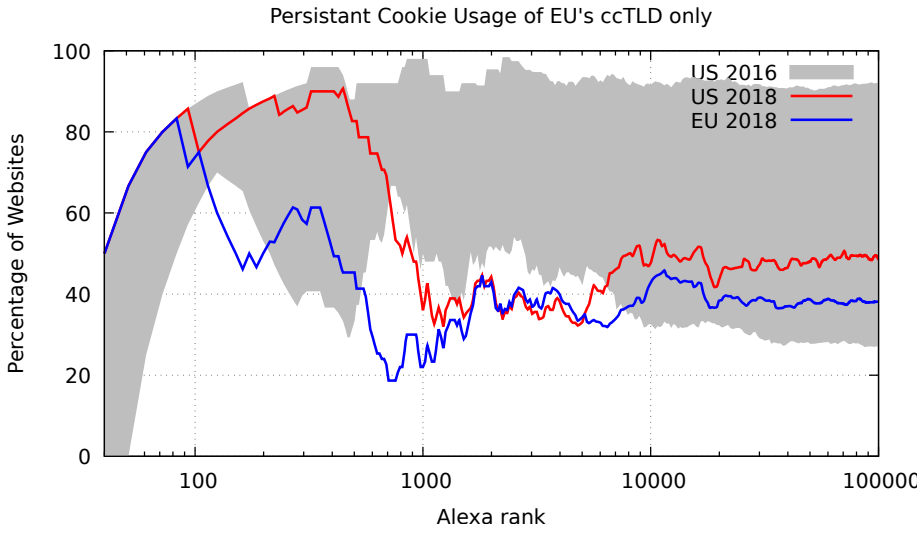

Fig. 4. Persistent cookie usage amount sites with EU-related domains, otherwise similar to Figure 1 and 3.

and present in all three data sets (Figure 4). Albeit some inaccuracies (e.g., vanity domains such as `start.at`, or multinationals such as `siemens.com`), the majority of those websites target EU consumers of some sort. 36% use cookies regardless of the geographic location of an visitor. 11% set cookies for U.S.-based but not EU users and 2% vice-versa. Between 31% and 46% of websites with EU-related domain names chose to drop cookie usage in the last two years, whereas 3% introduced them.

6 Impact on Measurement Studies

The adoption of the GDPR in the European Union has apparently created a two-class Internet with regard to privacy. While Internet users benefit from this development, it poses several challenges for Internet measurement studies. Future measurements on website behavior, cookie usage, privacy options, and related matters will heavily depend on the location of the measurement and provide potentially vastly different results. This diversification will further proceed should more world-regions (state unions, political super-powers) put effective Internet-regulation for privacy protection in place.

This implies reduced significance and universality, especially for measurements on privacy. Researchers will either have to restrict their scope to certain geographical regions or have to invest in multiple measurement sites repeating tests from each site. Additionally, as many websites present geographically-selected GDPR-compliance banners, measurement studies comparing the visual

representation of websites might lead to false results. Studies investigating cookies, e.g., the extent of tracking after providing consent, have to find a way to automatically overcome the banner to proceed to cookie setting.

But also beyond cookies, more challenges are caused by the adoption of the GDPR for measurement studies. For example, various European registrars no longer collect information for ICANN's WHOIS database as this line of action opposes GDPR's principle of data minimization. This impacts the database's usefulness for measurement studies as it limits the view on domain registrations in comparison to the era before GDPR adoption.

7 Limitations and Future Work

This study only compared snapshots of cookie usage by websites directly without third-party cookies (c.f. Iordanou et al. [17]). A stronger link between GDPR and changes in cookie usage would be possible by a longitudinal study during GDPR's adoption. Additionally, the 2016 data was collected with a slightly different methodology prohibiting to distinguish the temporal configuration of Javascript-set cookies, thus leading to upper and lower bounds for persistent cookies. Additionally, the different browsers might have triggered slightly different responses from websites.

Nevertheless, the current data suggests that EU's GDPR had a massive international impact on cookie usage in the web's ecosystem.

Some open topics remain for further investigation. First, we focused on persistent cookies in this study but neglected the use of sessions cookies which should become part of future inspections. Furthermore, as has been shown in the past, trackers can also rely on other information to identify users, like fingerprinting and local storage. Therefore the impact of GDPR on such types of user tracking needs to be analyzed as well. Additionally, since sites now need to ask the user's consensus before setting cookies, systems to detect and interact with such checks need to be devised to allow for a better analysis in future privacy measurements. Many cookies might still be set, after the user consent – however, if the user is willing to live with the banner, they might now enjoy many more websites cookie-free. We neglected, that some non-GDPR-compliant websites might offer the user to opt-out afterwards. Beyond, it should be investigated whether a website's real behavior is compliant with the declaration offered in the banners.

In our study we assumed geographical discrimination to be Source-IP or DNS-based. Other techniques exist, such as geographical routing announcements to different data centers.

8 Conclusion

General Data Protection Regulation (GDPR) has created a two-class Internet with regard to privacy. EU consumers encounter significantly less unconditional usage of persistent cookies when surfing the web than US visitors. 49.3% of cookie-using websites of the Alexa Top 1,000 choose to refrain from cookie setting

without consent on the first visit when facing an EU visitor, when they would for other visitors. This figure drops to an overall of 26% when observing all Alexa Top 100,000 websites. Further, the new regulations reduced cookie burdens for the rest of the Internet as well, i.e., even users from outside of the EU benefit from the GDPR's adoption and experience less cookies. When compared to data from 2016, the overall cookie load reduced by up to 46.7% for US consumers, albeit mostly for lower ranked websites.

For consumers, this is clearly great news for their privacy. However, researchers face an increasingly divided World Wide Web. Future Internet studies will have to account for different geographical regions or alternatively reduce their scope.

9 Acknowledgments

This research was funded by the *Christian Doppler Laboratory for Security and Quality Improvement in the Production System Lifecycle (CDL-SQI), Institute of Information Systems Engineering*, TU Wien and the Josef Ressel Centers project TARGET. The competence center SBA Research (SBA-K1) is funded within the framework of COMET Competence Centers for Excellent Technologies by BMVIT, BMDW, and the federal state of Vienna.

References

1. Amazon: Alexa top sites (2018), `https://aws.amazon.com/alexa/`, direct download: http://s3.amazonaws.com/alexa-static/top-1m.csv.zip
2. Barth, A.: HTTP State Management Mechanism. RFC 6265 (Proposed Standard) (Apr 2011). https://doi.org/10.17487/RFC6265, `https://www.rfc-editor.org/rfc/rfc6265.txt`
3. Cahn, A., Alfeld, S., Barford, P., Muthukrishnan, S.: An empirical study of web cookies. In: Proceedings of the 25th International Conference on World Wide Web. pp. 891–901. WWW '16, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland (2016). https://doi.org/10.1145/2872427.2882991
4. Cisco: Cisco umbrella 1 million (Dec 2016), `http://s3-us-west-1.amazonaws.com/umbrella-static/index.html`
5. Dabrowski, A., Merzdovnik, G., Kommenda, N., Weippl, E.: Browser history stealing with captive Wi-Fi portals. In: 2016 IEEE Security and Privacy Workshops (SPW). pp. 234–240 (May 2016). https://doi.org/10.1109/SPW.2016.42
6. Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., Holz, T.: We value your privacy ... now take some cookies: Measuring the GDPR's impact on web privacy. In: Network and Distributed System Security Symposium (NDSS) (2019)
7. Englehardt, S., Narayanan, A.: Online tracking: A 1-million-site measurement and analysis. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 1388–1401. CCS '16, ACM, New York, NY, USA (2016). https://doi.org/10.1145/2976749.2978313

8. Englehardt, S., Reisman, D., Eubank, C., Zimmerman, P., Mayer, J., Narayanan, A., Felten, E.W.: Cookies that give you away: The surveillance implications of web tracking. In: Proceedings of the 24th International Conference on World Wide Web. pp. 289–299. WWW '15, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland (2015). https://doi.org/10.1145/2736277.2741679
9. European Commission: Adequacy of the protection of personal data in non-EU countries (2018), https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en
10. European Commission: Cookies (2018), http://ec.europa.eu/ipg/basics/legal/cookies/index_en.htm
11. European Commission: Data transfers outside the EU (2018), https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu_en
12. Fielding, R., Reschke, J.: Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing. RFC 7230 (Proposed Standard) (Jun 2014). https://doi.org/10.17487/RFC7230, https://www.rfc-editor.org/rfc/rfc7230.txt
13. GDPR Shield: Gdpr shield (2018), https://gdprshield.co.uk
14. Gonzalez, R., Jiang, L., Ahmed, M., Marciel, M., Cuevas, R., Metwalley, H., Niccolini, S.: The cookie recipe: Untangling the use of cookies in the wild. In: 2017 Network Traffic Measurement and Analysis Conference (TMA). pp. 1–9 (June 2017)
15. Hannak, A., Soeller, G., Lazer, D., Mislove, A., Wilson, C.: Measuring price discrimination and steering on e-commerce web sites. In: Proceedings of the 2014 Conference on Internet Measurement Conference. pp. 305–318. IMC '14, ACM, New York, NY, USA (2014). https://doi.org/10.1145/2663716.2663744
16. Hern, A., Belam, M.: La times among us-based news sites blocking EU users due to GDPR (2018), https://www.theguardian.com/technology/2018/may/25/gdpr-us-based-news-websites-eu-internet-users-la-times
17. Iordanou, C., Smaragdakis, G., Poese, I., Laoutaris, N.: Tracing Cross Border Web Tracking. In: Proceedings of ACM IMC 2018. Boston, MA (October 2018)
18. Kulyk, O., Hilt, A., Gerber, N., Volkamer, M.: "this website uses cookies": Users' perceptions and reactions to the cookie disclaimer. In: European Workshop on Usable Security (EuroUSEC) (2018)
19. Lerner, A., Simpson, A.K., Kohno, T., Roesner, F.: 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 (2016), https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/lerner
20. Linden, T., Harkous, H., Fawaz, K.: The privacy policy landscape after the GDPR (2018), http://arxiv.org/abs/1809.08396
21. Merzdovnik, G., Huber, M., Buhov, D., Nikiforakis, N., Neuner, S., Schmiedecker, M., Weippl, E.: Block me if you can: A large-scale study of tracker-blocking tools. In: Security and Privacy (EuroS&P), 2017 IEEE European Symposium on. pp. 319–333. IEEE (2017)
22. Plonka, D., Berger, A.W.: kip: a measured approach to IPv6 address anonymization (2017), http://arxiv.org/abs/1707.03900
23. Pochat, V.L., van Goethem, T., Joosen, W.: Rigging research results by manipulating top websites rankings (2018), https://arxiv.org/abs/1806.01156v1

24. Scheitle, Q., Hohlfeld, O., Gamba, J., Jelten, J., Zimmermann, T., Strowes, S.D., Vallina-Rodriguez, N.: A long way to the top: Significance, structure, and stability of internet top lists. In: Proceedings of the Internet Measurement Conference 2018. pp. 478–493. IMC '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3278532.3278574
25. Sivakorn, S., Polakis, I., Keromytis, A.D.: The cracked cookie jar: HTTP cookie hijacking and the exposure of private information. In: 2016 IEEE Symposium on Security and Privacy (SP). pp. 724–742 (May 2016)
26. Tiku, N.: Europe's new privacy law will change the web, and more (2018), `https://www.wired.com/story/europes-new-privacy-law-will-change-the-web-and-more/`
27. Trammell, B., Kühlewind, M.: Revisiting the privacy implications of two-way internet latency data. In: Beverly, R., Smaragdakis, G., Feldmann, A. (eds.) Passive and Active Measurement. pp. 73–84. Springer International Publishing, Cham (2018)