Ensuring User Privacy in the Digital Age: A Quality-Centric Approach to Tracking and Data Protection

Vitalijs Teze[a] and Erika Nazaruka[b]
Faculty of Computer Science, Information Technology and Energy, Riga Technical University, 10 Zunda Embarkment, Riga, Latvia

Keywords: Digital Privacy, User Tracking, Data Protection, Privacy Risk Assessment, GDPR Compliance, Web Analytics, Online Security, Privacy-Focused Software.

Abstract: In the contemporary digital landscape, the pervasive practice of user tracking and the consequent erosion of data protection present significant challenges to user privacy. This paper introduces 'Privacy Risk Assessor' a software tool designed to evaluate and enhance online privacy. Addressing the dynamics of user tracking, the tool analyses websites for privacy-threatening metrics in the context of existing tracking mechanisms. Employing a methodological approach, the tool's architecture enables efficient processing and adaptability to tracking techniques and privacy regulations. The research focuses on key metrics of quality attributes including security, usability, trust, reliability, and performance, providing actionable insights into privacy risks. An evaluation was conducted on a dataset of 492 Latvian websites, with an emphasis on diverse privacy-related factors. The study revealed insights into prevalent privacy practices and underscored the tool's effectiveness in real-world scenarios. The 'Privacy Risk Assessor' stands out for its possibility to be integrated into web development process, offering developers and end-users ability to proactively measure potential privacy threats.

1 INTRODUCTION

The digital age presents both vast opportunities and significant privacy challenges, particularly through browser fingerprinting and tracking technologies that threaten privacy. This paper explores how advertising-driven data collection, especially the transition to behavioural advertising, commodifies user data without clear consent, monitoring user behaviour in a complex ecosystem.

In response to these challenges, we indicate the existence of legislative landscape, notably the General Data Protection Regulation (GDPR) (European Parliament, 2018) in the European Union, designed to empower users and safeguard privacy. Additionally, the ePrivacy Directive (2002/58/EC) (European Parliament, 2002), also known as the "Cookie Law," amended by Directive 2009/136/EC (European Parliament, 2009), explicitly addresses cookies, and requires consent for storing or accessing information on a user's device. However, the existing legal frameworks, while providing a basis for recourse, fall short of preventing data misuse.

This backdrop sets the stage for the introduction of 'Privacy Risk Assessor' (hereinafter, PRA), a tool developed to tackle these privacy challenges head-on. Unlike reactive privacy tools currently in the market, PRA offers a proactive, automated approach to privacy protection. It is designed to dynamically measure privacy threats and provide statistics of potentially threatening behaviours a given sites poses to its user.

This paper outlines the PRA's architecture, its adaptability to different web environments, and its role in assessing privacy risks across 492 Latvian websites. It contributes to digital privacy puzzle by offering both theoretical perspectives and practical tools to address contemporary privacy challenges.

Our main contributions include proposing privacy evaluation metrics for analyzing potential website privacy threats, developing a tool that visits and calculates these metrics for a set of websites for

[a] https://orcid.org/0009-0001-2165-4789
[b] https://orcid.org/0000-0002-1731-989X

Teze, V. and Nazaruka, E.
Ensuring User Privacy in the Digital Age: A Quality-Centric Approach to Tracking and Data Protection.
DOI: 10.5220/0012565800003687

In *Proceedings of the 19th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2024)*, pages 299-306
ISBN: 978-989-758-696-5; ISSN: 2184-4895

further evaluation, which we share with those interested in contributing [1], and performing data collection and analysis through an experiment that measures request data from Latvian websites.

Section 2 of this paper delves into the evolution of digital privacy concerns, existing solutions, and legislative measures like GDPR, identifying gaps in current approaches to digital privacy.

In Section 3 we describe the architecture of the PRA explaining its modular design and integration capabilities.

Section 4 outlines several key metrics developed to assess privacy risks on websites. These metrics include the analysis of third-party domains, cookies, and requests to third-party domains, among others.

Section 5 details the methodology for calculating the proposed metrics, describing the process of data collection and analysis

Section 6 describes the practical application of PRA in evaluating 492 Latvian websites. It discusses the findings from this evaluation, showcasing the tool's ability to effectively analyze and report on digital privacy metrics.

Section 7 evaluates the effectiveness of our developed metrics in identifying privacy risks and the broader significance of these results for enhancing digital privacy protections and legislative compliance.

Section 8 highlights the contributions of the PRA to the field and suggests avenues for future research and development in digital privacy protection.

2 BACKGROUND AND RELATED WORK

The digital era has brought significant challenges to user privacy and data protection. The issue at hand is the pervasive use of browser fingerprinting and user tracking technologies that pose a substantial threat to individual privacy. With every online interaction, users potentially expose personal information to various entities, ranging from benign data collectors to malicious actors. The advent of sophisticated tracking mechanisms, such as cookies, web beacons, and fingerprinting scripts, has made it exceedingly difficult for users to maintain anonymity and control over their personal data (Boda et al., 2012; Elbanna & Abdelbaki, 2018; Laperdrix et al., 2019; Nikiforakis et al., 2013).

In the realm of web interactions, cookies serve as a fundamental mechanism to preserve the state across otherwise stateless HTTP sessions(Kristol & Montulli, 1997), thereby enhancing user navigational experiences. GDPR represents a paradigm shift in the approach to web privacy, mandating that explicit consent be obtained for the storage of cookies not indispensable to a website's core functionalities.

Advertising has been one of the primary drivers for the development and deployment of tracking technologies. The transition from traditional advertising to behavioral targeting has led to a complex ecosystem where user data is a valuable commodity (Beales, 2010; Provost et al., 2009; Wu et al., 2009; Yan et al., 2009). The ability to deliver personalized advertisements hinges on the extensive collection and analysis of user behavior, often without explicit consent(Confessore, 2018; Dehling et al., 2019; Ur et al., 2012).

Despite the benefits of personalized content, the implications for user privacy are profound. Research has demonstrated that even seemingly innocuous data can, over time, be aggregated to construct detailed and invasive user profiles (Estrada-Jiménez et al., 2017; Woensdregt et al., 2019). These profiles can then be exploited for various purposes, not all of which align with the users' best interests.

The introduction of privacy legislation, such as GDPR in the EU, represents a significant step towards empowering users and protecting privacy (European Parliament, 2018; Porter, 2019). However, while these regulations offer a framework for legal recourse, they do not inherently prevent the collection or misuse of data.

Efforts to combat these privacy issues have led to the development of a range of solutions. From ad blockers and privacy-focused browsers to anti-tracking extensions, these tools aim to give users more control over their data (Bennett, 2019; Brave, 2019; Brave Browser, n.d.). Yet, none provide complete privacy protection. The limitations of these tools often stem from the reactive nature of their development; as new tracking methods emerge, privacy tools must adapt accordingly(Fishback, 2019; Kristol & Montulli, 1997; Sullivan, 2018).

Recent studies have unveiled advanced tracking techniques that pose additional challenges to user privacy, notably through Domain Name System (DNS) CNAME redirections. This method, known as CNAME cloaking, enables third-party trackers to masquerade as first-party entities, thereby bypassing traditional cookie policies and ad-blockers. Such practices have been identified to facilitate the exfiltration of sensitive cookies, including those

[1] https://github.com/vitally/Privacy-Risk-Assessor

carrying authentication information, to third-party domains without user consent. This discovery underscores the limitations of current browser and regulatory defenses, emphasizing the need for more comprehensive solutions to protect against such invasive tracking methods (Ren et al., 2021).

Another significant development that complicates the digital privacy landscape is the innovative use of first-party cookies for tracking (Chen et al., 2021) that involves third-party scripts executing code to directly set cookies without utilizing the conventional "Set-Cookie" headers, underscoring the dynamic nature of tracking strategies, adapting to navigate the evolving web privacy frameworks.

The need for a proactive and automated approach to privacy protection is clear. This is where the software tool proposed in this research comes into play. Unlike existing solutions, this tool is designed to measure privacy threats in a more dynamic and user-centric manner. It employs a strategy that measures potential tracking threats and calculates metrics allowing to make cumulative evaluation.

3 PRIVACY RISK ASSESSOR HIGH-LEVEL ARCHITECTURE

Within the website development lifecycle, privacy often receives insufficient attention, risking GDPR non-compliance and associated penalties. The PRA tool, primarily aimed at enterprise use, facilitates automated privacy threat analysis to heighten awareness among developers and users. Its modular architecture ensures adaptability to new tracking technologies and legal requirements, making it ideal for both extensive and specific website evaluations. Integrated into CI/CD pipelines, it offers continuous privacy monitoring. Developed with Node.js for efficient asynchronous processing and MongoDB for flexible data management, it streamlines privacy assessments across diverse data types.

The PRA software employs a modular architecture for scalability and efficient processing, consisting of interlinked components for various functions: a Server Module for API call handling, a Database Client Module for database interactions, a Database Helper for data manipulation, an API Module for request processing, a Worker Factory for asynchronous tasks, Site Retriever and Helper Modules for website listing, a Site Visitor Module for orchestrating site visits, a Navigation Module using Puppeteer for real user simulation, an Analysis Module for data processing and privacy risk evaluation, a URL Module for URL tasks, a WHOIS(Daigle, 2004) Module for site ownership information, and a Configuration Module for system settings management. Module interdependencies are shown on Figure 1.

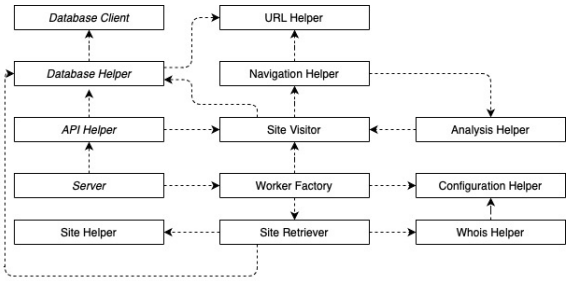

Figure 1: Modular architecture of PRA.

4 PROPOSED QUALITY METRICS

Our proposed tool employs passive observation to study initial user-website interactions without direct engagement, underlining the importance of GDPR-compliant essential cookies for website functionality. This approach forms the basis for evaluating websites' adherence to GDPR consent requirements. We suggest that transmitted data volume correlates with privacy risk, as each bit's addition exponentially increases potential data uniqueness, highlighting the importance of scrutinizing website-third-party exchanges that elevate privacy risks. Analysing these interactions, especially involving third-party cookies, is key to distinguishing between necessary and excessive cookie use, directly affecting privacy compliance assessment. We observe that first-party websites may engage third parties without needing to maintain session state via cookies, questioning the necessity of third-party cookies for functionality versus tracking. The pattern of data exchange suggesting user tracking upon initial website visit underscores privacy concerns, necessitating thorough investigation into data collection and usage purposes.

Thus, focusing on cookie analysis, particularly third-party cookies, aims to differentiate essential from non-essential cookie uses to uphold user privacy. The study prioritizes metrics related to third-party domain interactions, offering insights into data exchanges that may compromise privacy by enabling user profiling and targeted communication. This approach informs our scrutiny of privacy preservation practices and regulatory compliance.

Each metric within the framework is calibrated such that a value of zero denotes the absence of detectable privacy threats, reflecting a baseline of optimal privacy conditions. Conversely, incrementally higher values are indicative of escalating privacy risks. This scaling ensures that the metrics serve as quantitative indicators of potential privacy infringements.

Number of Distinct Third-Party Domains Addressed (**M01**): This metric quantifies the distinct third-party domains interacted with by the website. While certain requests to these domains may be benign, such as retrieving fonts or images, others may involve more intrusive activities like tracking the user. These tracking requests might transmit browser data to third parties, potentially without the user's explicit consent. The distinction between harmless and privacy-invasive requests underlines the importance of this metric in evaluating a website's approach to user privacy.

Number of Distinct Third-Party Cookie Domains (**M02**): This metric quantifies the unique third-party domains referenced in the 'domain' property of cookies. It assesses the diversity of external entities interacting with the user via cookies, providing insight into the breadth of third-party engagement and potential data sharing channels.

Number of Third-Party Cookies (**M03**): This metric enumerates cookies whose 'domain' attribute differs from the visited site's domain. It is crucial for evaluating the extent of third-party tracking activities embedded within a site, highlighting the presence of external entities that may be collecting user data.

Number of Cookies Set by Third-Party Requests (**M04**): This metric tallies the cookies set in response to third-party requests. It provides a measure of the direct impact of third-party interactions on the user's browser, indicating the level of third-party influence in terms of cookie-based tracking and data storage.

Number of Third-Party Frames Addressed (**M05**): This metric reflects the count of third-party requests made via usage of frames build into the website's markup.

Covert Third-Party Cookie Number (**M06**): This metric assesses the extent to which cookies, typically associated with third-party services, are set under domains that match the visited site, potentially masquerading as first-party cookies. This is important for evaluating the transparency of cookie usage and for identifying tracking practices that may not be clear to users.

Number of Cookies Set via CNAME Cloaking (**M07**): This metric counts the number of cookies where the cookie's domain is different from the domain resulting from any CNAME redirection applied to the request that set the cookie. It aims to identify tracking mechanisms that exploit DNS CNAME redirections to bypass conventional third-party cookie policies, highlighting a sophisticated method of tracking that masquerades third-party requests as first-party. This metric is crucial for uncovering covert tracking practices that may not be immediately apparent, further enhancing the tool's capability to assess privacy risks associated with modern web tracking techniques.

Domain Transparency (**M08**): This binary metric evaluates the clarity of domain ownership based on WHOIS lookup results. When the owner is a recognizable company, it implies a commitment to transparency. Conversely, domains registered to private individuals with GDPR-masked data or unidentifiable ownership may suggest an intentional effort to obscure domain ownership. The inability to identify the domain owner results in the flag being raised.

User Consent Compliance (**M09**)**:** This binary metric assesses data collection before explicit user consent, focusing on third-party cookies being set by requests made at the initial site visit. Activity of potential data collection without consent signals non-compliance with consent norms, triggering the metric's flag.

Third-Party Request Diversity Index (**M10**): This metric employs the Shannon Diversity Index (Kiernan, 2023)calculated by the Formula 1,

$$H = -\Sigma p_i * \ln(p_i) \quad (1)$$

commonly used in ecological studies, to analyze the diversity of third-party domains addressed by a website. Each domain's frequency of requests contributes to the index, providing a quantifiable measure of diversity. A higher index value indicates a more varied range of third-party requests, potentially signaling a greater likelihood of unintended user data exposure. This metric facilitates comparative analysis of the third-party interaction landscape across different websites.

Cookies Set by Third-Party Requests Diversity Index (**M11**): Adopting the Shannon Diversity Index, this metric is applied to evaluate the variety of cookies set in response to third-party requests, based on the association of each cookie to its originating third-party domain. The diversity index here quantifies the range of different entities setting cookies during a site visit. A higher diversity index suggests a broader spectrum of cookies from various domains being stored, which can be indicative of extensive third-party involvement and varied data

collection practices. This metric is instrumental for comparing cookie-setting behaviors across multiple sites.

Third-Party Request Execution Time (**M12**): This metric aggregates the cumulative execution time, measured in milliseconds, of all third-party requests initiated by a website. While many of these requests may operate asynchronously and thus do not directly contribute to perceived loading delays, the total execution time serves as an approximation of the resources allocated to activities that may involve user tracking. By quantifying the time spent on third-party requests, this metric provides an insight into the extent of external interactions and their potential impact on both website performance and privacy implications.

While no single metric conclusively indicates tracking, analyzing them collectively offers a comprehensive view of a website's potential privacy threats. This holistic approach enables a comparative analysis between websites with divergent behavioral patterns. For instance, a website characterized by an absence of third-party requests and cookies, coupled with transparent domain ownership, would inherently be deemed less intrusive from a privacy perspective. In contrast, a website engaging in multiple requests, a portion of which culminates in the setting of cookies through CNAME Cloaking, warrants scrutiny. Such behavior suggests a greater potential of utilizing a sophisticated approach to user tracking, meriting further investigation. This dichotomy underscores the importance of a composite metric analysis in identifying websites that pose significant privacy risks. Subsequently, sites exhibiting concerning patterns should be subjected to an in-depth analysis, either manual or automated, to elucidate the nature and extent of potential privacy infringements.

Summarized metric evaluation matrix is presented in Table 1.

Table 1: Privacy Risk Metric Evaluation Values.

Metric	No-Risk Value	Elevated Risk Value
M01-M07, M10-M12	0	> 0
M08, M09	0	1

5 METRIC CALCULATION METHODOLOGY

Recording Third-Party Requests: All outbound requests to third-party domains are recorded and stored. This facilitates the computation of the 'Number of Distinct Third-Party Domains Addressed' metric by enumerating unique second-level domain names targeted by these requests.

Measuring Response Times: The temporal difference in milliseconds between the initiation of a request and the receipt of its response is calculated. This enables the derivation of the 'Third-Party Request Execution Time'.

Mapping and Analyzing Request Diversity: Each third-party request is cataloged in a map keyed by the second-level domain name, incrementing a corresponding integer count per request. Applying Shannon's Diversity Index formula to this map facilitates the calculation of the 'Third-Party Request Diversity Index'.

Assessing Cookie Setting Behaviors: Analysis of responses for 'set-cookie' headers and subsequent storage of identified cookies aids in determining the 'Number of Cookies Set by Third-Party Requests' metric.

Counting Total Number of Third-Party Cookies: The aggregation of the 'Number of Cookies Set by Third-Party Requests' metric with the count of third-party cookies obtained via the Chrome Devtools Protocol (Google, 2023) yields the 'Number of Third-Party Cookies' metric.

Evaluating Third-Party Cookie Domain Diversity: The creation of a map correlating second-level domain names in cookies' 'domain' property to their quantity enables the computation of both 'Number of Distinct Third-Party Cookie Domains' and 'Cookies Set by Third-Party Requests Diversity Index' metrics.

Third-Party Frame Analysis: For sites containing multiple frames, those addressing URLs with second-level domains divergent from the visited site contribute to the 'Number of Third-Party Frames Addressed' metric.

Assessing User Consent Compliance: Given the software's single, non-interactive site visitation approach, the presence of third-party cookies at site load without user consent yields a '1' for the 'User Consent Compliance' metric, while their absence results in a '0'.

Identifying Covert Third-Party Cookies: The system maintains a record of cookie names associated with common advertising or tracking services (e.g., '_ga' from Google Analytics, '_fbp' from Meta's Facebook Pixel). Cookies matching these names with a 'domain' property aligning with the visited site's domain are classified as 'Covert Third-Party Cookies'.

Table 2: Metric Descriptive Statistics.

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
Median	7	0	0	0	0	0	0	0	0	1.466	0	3795
Mean	8.614	0.551	1.24	1.268	0.763	0.783	2.071	0.162	0.301	1.424	0.108	8889.99
Std. Deviation	7.08	1.088	3.602	3.609	1.319	1.308	15.225	0.369	0.459	0.772	0.314	22101.263
Variance	50.121	1.185	12.978	13.027	1.741	1.71	231.808	0.136	0.211	0.597	0.099	$4.885\times10^{+8}$
Minimum	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	49	8	45	45	10	6	153	1	1	3.19	1.991	273491

Identifying cookies that use CNAME Cloaking: Each cookie-setting request from sites is analysed by conducting a DNS query to check for CNAME redirection. If found, we compare the redirection's target domain with the cookie's domain attribute. Discrepancies suggest CNAME Cloaking, enabling accurate identification and counting of such cookies.

6 TOOL EVALUATION ON A SET OF LATVIAN SITES

To evaluate the capabilities of the PRA tool, a set of 492 Latvian websites was compiled for analysis. This dataset was sourced from Netcraft's compilation of top sites in Latvia[4] and included all sites within the '.lv' domain zone listed in BuiltWith's top 1 million sites[5].

Table 2 shows descriptive statistics for metrics from Section 4, calculated as per Section 5 methodologies after site evaluations. These metrics are aggregated into a correlation heatmap in Figure 2, offering key insights into potential privacy threats through their interrelations.

Firstly, there are positive intuitive correlations among several pairs of metrics, notably between the number of third-party cookies (M03) and the number of cookies set by third-party requests (M04).

Additionally, the number of distinct third-party domains addressed (M01) shows positive correlations with several other metrics, suggesting that sites interacting with more third-party domains are likely to have higher instances of third-party cookies and potentially a more diverse range of third-party requests.

Furthermore, the execution time of third-party requests (M12) displays varying degrees of correlation with other metrics. This variation indicates that while the execution time is influenced by the number and diversity of third-party requests, it is not solely dependent on them.

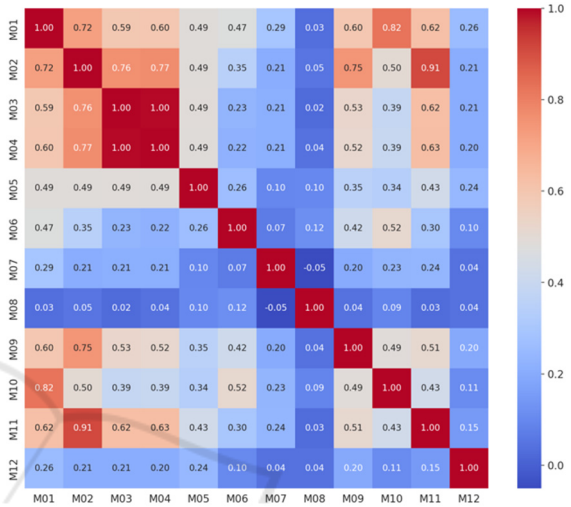

Figure 2: Collected metric correlation heatmap.

7 DISCUSSION

In the interpretation of results, our analysis of privacy metrics across 492 Latvian websites reveals insights into third-party domain interactions, cookie usage, and tracking mechanisms. Metrics M01 and M03 show that most sites engage with a limited number of third-party domains and set few third-party cookies, suggesting a general tendency towards minimizing external dependencies. However, the presence of outliers indicates that a subset of websites extensively uses third-party cookies and domains, potentially increasing privacy risks.

The skewness observed in metrics M02, M04, and M05 towards lower values, with notable exceptions, highlights a concentration of websites that either limit third-party engagements or employ a more consolidated approach to user tracking. This could reflect a cautious stance on privacy or a focused use of third-party resources. The variability in the presence of covert third-party cookies (M06) and the use of CNAME cloaking (M07) points to sophisticated tracking techniques on some sites,

[4] https://trends.netcraft.com/topsites?c=LV

[5] https://builtwith.com/top-1m

emphasizing the need for advanced detection methods.

Domain transparency (M08) and user consent compliance (M09) serve as binary indicators of websites' commitment to privacy norms, with our findings suggesting a mixed landscape of adherence. The diversity indices (M10 and M11) and third-party request execution time (M12) further differentiate websites, showing a broad spectrum of practices from minimal to extensive third-party integrations, which could impact both privacy and user experience.

Normalization of metrics underscores the relative privacy threat of each site, with la.lv, tavacena.lv, and azeta.lv identified as the top three posing the highest potential risks.

Our tool's methodology aimed to detect cookie setting via JavaScript and tracking with the Canvas HTML element, yet found no such instances among the websites analyzed. This outcome prompts a critical examination of whether the lack of detected tracking is due to our method's limitations or truly reflects website practices. This uncertainty emphasizes the need for further research to refine detection techniques or verify these initial findings, marking a crucial direction for future work in digital privacy assessment.

The data collection process encompassed the comprehensive capture of request parameters, request bodies, and cookie expiration dates. Future iterations of the research and the continued development of the tool will explore the analysis of payload content and the identification of data flows. This advancement aims to deepen our understanding of how data is managed and transferred across sites, offering further insights into privacy practices and potential vulnerabilities.

The tool from our research can be integrated into corporate development and audit processes, allowing organizations to assess websites for privacy risks. This is key in preventing data leaks by identifying threats from employee internet usage, thereby strengthening an organization's cybersecurity, and protecting sensitive data from unauthorized exposure.

8 CONCLUSIONS

In conclusion, our study presents a comprehensive privacy risk assessment across 492 Latvian websites, utilizing the Privacy Risk Assessor tool to analyze and quantify privacy threats through various metrics. The findings reveal a spectrum of practices regarding third-party engagements, cookie usage, and tracking mechanisms, highlighting the importance of robust privacy protection measures. This research underscores the need for ongoing vigilance and refinement of privacy assessment tools to address sophisticated tracking techniques and comply with evolving privacy regulations.

Further investigation is warranted to enhance the tool's detection capabilities, particularly in identifying JavaScript-based cookie setting and Canvas element tracking. The absence of these practices among the evaluated websites poses questions about measurement methodologies or the actual prevalence of such tracking techniques, pointing towards areas for future research.

The potential integration of the Privacy Risk Assessor into enterprise development and audit processes suggests a proactive approach to safeguarding digital privacy and preventing corporate data leaks. By enabling companies to assess the privacy risks of websites accessible to their employees, this tool contributes significantly to the broader discourse on digital privacy and security, advocating for a more transparent, consent-based online environment.

REFERENCES

Beales, H. (2010). The Value of Behavioral Targeting. *Director*, 1–23. http://www.socialized.fr/wp-content/uploads/2010/08/beales-etude-sur-le-ciblage-comportemental-sur-internet-ppc4bible.pdf

Bennett, C. (2019). *Behind the One-Way Mirror: A Deep Dive Into the Technology of Corporate Surveillance*. https://www.eff.org/wp/behind-the-one-way-mirror

Boda, K., Földes, Á. M., Gulyás, G. G., & Imre, S. (2012). User Tracking on the Web via Cross-Browser Fingerprinting. In *LNCS* (Vol. 7161, pp. 31–46). https://doi.org/10.1007/978-3-642-29615-4_4

Brave. (2019). *Brave Launches the First Advertising Platform Built on Privacy*. https://brave.com/brave-ads-launch/

Brave Browser. (n.d.). Retrieved November 13, 2020, from https://brave.com/

Chen, Q., Ilia, P., Polychronakis, M., & Kapravelos, A. (2021). Cookie Swap Party: Abusing First-Party Cookies for Web Tracking. *Proceedings of the Web Conference 2021*, 2117–2129. https://doi.org/10.1145/3442381.3449837

Confessore, N. (2018, April 4). *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*. https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Daigle, L. (2004, September). *WHOIS Protocol Specification*. https://datatracker.ietf.org/doc/html/rfc3912

Dehling, T., Zhang, Y., & Sunyaev, A. (2019). Consumer perceptions of online behavioral advertising. *Proceedings - 21st IEEE Conference on Business*

Informatics, CBI 2019, *1*, 345–354. https://doi.org/10.1109/CBI.2019.00046

Elbanna, A., & Abdelbaki, N. (2018). Browsers fingerprinting motives, methods, and countermeasures. *CITS 2018 - 2018 International Conference on Computer, Information and Telecommunication Systems*. https://doi.org/10.1109/CITS.2018.8440163

Estrada-Jiménez, J., Parra-Arnau, J., Rodríguez-Hoyos, A., & Forné, J. (2017). Online advertising: Analysis of privacy threats and protection approaches. *Computer Communications*, *100*, 32–51. https://doi.org/10.1016/j.comcom.2016.12.016

European Parliament. (2002, July 12). *Directive 2002/58/EC of the European Parliament and of the Council*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32002L0058&qid=1707565293441

European Parliament. (2009, November 25). *Directive 2009/136/EC of The European Parliament And Of The Council*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32009L0136&qid=1707565614743

European Parliament. (2018, May 25). *General Data Protection Regulation*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504

Fishback, G. (2019). *Google's New Cookie Restrictions: A Nail in a Coffin We Should Have Already Buried*. https://www.martechadvisor.com/articles/ads/google-new-cookie-restrictions-a-nail-in-a-coffin-we-should-have-already-buried/

Google. (2023). *Chrome DevTools Protocol*. https://chromedevtools.github.io/devtools-protocol/

Kiernan, D. (2023). *NATURAL RESOURCES BIOMETRICS*. LibreTexts.

Kristol, D., & Montulli, L. (1997). *Request for Comments: 2109*. https://tools.ietf.org/html/rfc2109

Laperdrix, P., Bielova, N., Baudry, B., & Avoine, G. (2019). *Browser Fingerprinting: A survey*. http://arxiv.org/abs/1905.01051

Nikiforakis, N., Kapravelos, A., Joosen, W., Kruegel, C., Piessens, F., & Vigna, G. (2013). Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. *Proceedings - IEEE Symposium on Security and Privacy*, 541–555. https://doi.org/10.1109/SP.2013.43

Porter, J. (2019). *Google fined €50 million for GDPR violation in France*. https://www.theverge.com/2019/1/21/18191591/google-gdpr-fine-50- million-euros-data-consent-cnil

Provost, F., Dalessandro, B., & Hook, R. (2009). Audience selection for on-line brand advertising: Privacy-friendly social network targeting. *Proceedings of the Fifteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 707–715.

Ren, T., Wittmany, A., Carli, L. De, & Davidsony, D. (2021, August 15). An Analysis of First-Party Cookie Exfiltration due to CNAME Redirections. *Proceedings 2021 Workshop on Measurements, Attacks, and Defenses for the Web*. https://doi.org/10.14722/madweb.2021.23018

Sullivan, L. (2018). *64% Of Tracking Cookies Are Blocked, Deleted By Web Browsers*. https://www.mediapost.com/publications/article/316757/64-of-tracking-cookies-are-blocked-deleted-by-we.html

Ur, B., Leon, P. G., Cranor, L. Faith., Shay, R., & Wang, Y. (2012). Smart, useful, scary, creepy. *Proceedings of the Eighth Symposium on Usable Privacy and Security - SOUPS '12*, 1. https://doi.org/10.1145/2335356.2335362

Woensdregt, J. W., Al-Khateeb, H. M., Epiphaniou, G., & Jahankhani, H. (2019). AdPExT: Designing a Tool to Assess Information Gleaned from Browsers by Online Advertising Platforms. *Proceedings of 12th International Conference on Global Security, Safety and Sustainability, ICGS3 2019*. https://doi.org/10.1109/ICGS3.2019.8688328

Wu, X., Yan, J., Liu, N., Yan, S., Chen, Y., & Chen, Z. (2009). Probabilistic latent semantic user segmentation for behavioral targeted advertising. *Proceedings of the Third International Workshop on Data Mining and Audience Intelligence for Advertising - ADKDD '09*, 10–17. https://doi.org/10.1145/1592748.1592751

Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., & Chen, Z. (2009). How much can behavioral targeting help online advertising? *Proceedings of the 18th International Conference on World Wide Web - WWW '09*, 261. https://doi.org/10.1145/1526709.1526745