

Over-Reliance on Third-Party Cookies: How It May Backfire on Digital Advertising—Insights from the Cookie Dependency Index(CDI)

Chen-Yu Chen¹⁺, Chiang-Yu, Cheng²

^{1 2}Department of Data Science, Soochow University, Taiwan

Abstract. Third-party cookies play a critical role in tracking, privacy management, and digital advertising. However, increasing GDPR regulations and Privacy Sandbox initiatives are driving their phase-out, posing challenges for websites that are heavily reliant on these technologies. This study introduces the Cookie Dependency Index (CDI) to quantify third-party cookie reliance by analyzing data from 100 high-traffic e-commerce websites. Using a four-quadrant risk model and detailed case studies, the research identifies varying risks faced by different website types and proposes digital marketing strategies to mitigate these risks. The proposed framework serves as a quantitative risk assessment tool, offering practical insights to help the online advertising and e-commerce sectors adapt to evolving privacy regulations.

Keywords: cookies, third-party cookies, tracking, privacy, GDPR, online advertising, Privacy Sandbox.

1. Introduction

Cookies are essential for website operations, user experience, and digital advertising. Users often encounter consent banners, yet their functionality remains misunderstood. Invented in 1994, cookies were designed as a mechanism to maintain state between clients and servers [1]. In simple terms, cookies are small files or data stored on users' devices by websites, allowing them to "remember" specific user information, thereby enhancing personalization and interactivity.

Based on their domain of origin, cookies can be categorized into two types: first-party cookies and third-party cookies. First-party cookies are set by the website the user is currently visiting and are mainly used to store login statuses, preference settings, and personalized content. In contrast, third-party cookies are set by external domains not currently being visited and are often used for cross-site tracking, behavioral analysis, and targeted advertising. For instance, when a user views a particular product on an e-commerce site, they may later see related advertisements on other websites [2], illustrating third-party cookies' role in digital advertising.

In 2010, *The Wall Street Journal* conducted the groundbreaking "What They Know" investigation, revealing that over half of the top 50 U.S. had installed at least 23 third-party cookies, with some exceeding 100. A 2016 study found that 94% of the top 100,000 websites contained at least one cookie, up from 67% in 2009, illustrating the growing dominance of web tracking. By 2020, HTTP Archive reported that 84% of 5.8 million homepages used cookies, with 79% specifically relying on third-party cookies[3]. This highlights the widespread use of third-party cookies in data-driven advertising and user behavior analysis.

While cookies enhance user experience and targeted advertising, they also raise privacy concerns. Growing awareness has led to increased scrutiny of web tracking practices. This prompted regulatory bodies and browser developers to take action. In response, the European Union enacted GDPR in 2018, which has significantly impacted cookie regulations. To comply, many major companies revised their privacy policies. Furthermore, GDPR required websites to obtain explicit user consent for cookie usage[4].

GDPR has had a global impact on privacy regulations. However, compliance remains challenging, as any online service operating within Europe must comply with its regulations. A study of the top 10,000 UK websites found that only 11.8% met GDPR's requirements. Common violations included pre-checked consent boxes, hidden "reject" options, and continued tracking after opt-out [5]. These enforcement challenges suggest that stricter regulations may follow.

⁺ Corresponding author. Tel.: +886 935 673 608
E-mail address: anguschen10830@gmail.com

In 2020, Google took action in response to privacy regulations by announcing its plan to phase out third-party cookies in Chrome and introduce Privacy Sandbox as an alternative. This sparked extensive discussions about supporting a privacy-friendly advertising ecosystem. Originally set for 2023, this transition has been delayed multiple times and is now expected by 2025 [6,7]. Regardless, the elimination of third-party cookies is bound to have a significant impact on website data collection models and advertising strategies.

Although GDPR has been implemented, compliance among websites remains inconsistent, suggesting that future regulations may tighten restrictions. Google's uncertain timeline for cookie deprecation leaves website operators facing significant risks. In this shifting environment, once third-party cookies are discontinued, many website operators may face serious threats. The inability to track users will reduce ad targeting efficiency, weaken behavioral insights, and lower ad revenue, especially for ad-dependent industries[8]. In addition to the tracking difficulties and ad revenue declines resulting from the phase-out of third-party cookies, recent studies estimate that limiting cookie lifetimes to one year could reduce EU cookie-based display ad revenues by about 8.5%—equivalent to roughly €904 million in annual losses. This further highlights the dual challenges faced by the digital advertising ecosystem under increasingly stringent privacy regulations[21].

While many studies have focused on the privacy implications of third-party cookies, there remains a significant gap in the literature — a standardized and quantitative method to measure website dependency on these technologies. Understanding a website's reliance on third-party cookies is crucial for assessing risks associated with cookie deprecation. To address this gap, this paper introduces a novel metric, the Cookie Dependency Index (CDI), which assesses the extent of a website's reliance on third-party cookies. By utilizing Cookie Scanner technology, this study evaluates 100 high-traffic e-commerce websites to validate CDI's applicability and effectiveness. Through data analysis and visualization using a four-quadrant risk model and detailed case analyses, this study categorizes websites based on risk levels, illustrating specific challenges and adaptation strategies. Additionally, we examine Google's Privacy Sandbox and its role in shaping future privacy and data ecosystems, offering insights for advertisers, e-commerce businesses, and regulators.

2. Methodology

This study quantifies e-commerce websites' reliance on third-party cookies and evaluates the impact of their phase-out. We introduce the Cookie Dependency Index (CDI) as a standardized measure and apply data collection, quadrant analysis, and visualization techniques to assess risk levels.

2.1. Cookie Dependency Index (CDI)

Currently, there is no unified method for quantifying third-party cookie dependency. Thus, this study defines the Cookie Dependency Index (CDI) as:

$$CDI = \frac{\text{Third-party cookies}}{\text{First-party cookies} + \text{Third-party cookies}}, \text{ where CDI ranges from 0 to 1}$$

- A higher CDI indicates heavy reliance on third-party cookies, meaning the website will likely face significant tracking difficulties and revenue decline once third-party cookies are phased out.
- A lower CDI suggests that the website primarily relies on first-party data or alternative technologies, minimizing the impact of third-party cookie removal and reducing tracking difficulties.

2.2. Cookie Scanners

This study utilizes “Cookie-Scanner” and “Cookie Metrix” to collect first-party and third-party cookie data [9-10]. Initially, multiple scanning tools testing revealed variations in scanning accuracy due to technical implementation, update frequency, and parsing methods[11], consistent with prior research. Since automated tools differ in architecture, data inconsistencies may arise [12]. To enhance accuracy and reliability, the study adopted cross-verification using two tools, reducing potential biases from a single scanner.

2.3. Data Selection and Collection

The initial research design aimed to select the top 100 international e-commerce websites from SimilarWeb and Semrush (October 2024) as samples. However, some sites could not be successfully scanned due to technical constraints, leading to incomplete data. This aligns with past findings that dynamic content loading,

JavaScript rendering, and cookie consent banners hinder automated tools from capturing all cookies [12]. To ensure accuracy, the study employed a multi-tool cross-verification approach, using a secondary tool when the primary scanner failed to ensure data completeness and accuracy.

The final dataset refined the sample to 100 high-traffic international e-commerce websites, spanning retail, fashion, electronics, and household goods, with traffic ranging from 33.09M to 5450M visits. This study follows Figure 1 as the overall research process for sample selection, cookie scanning, CDI computation, and visualization analysis.

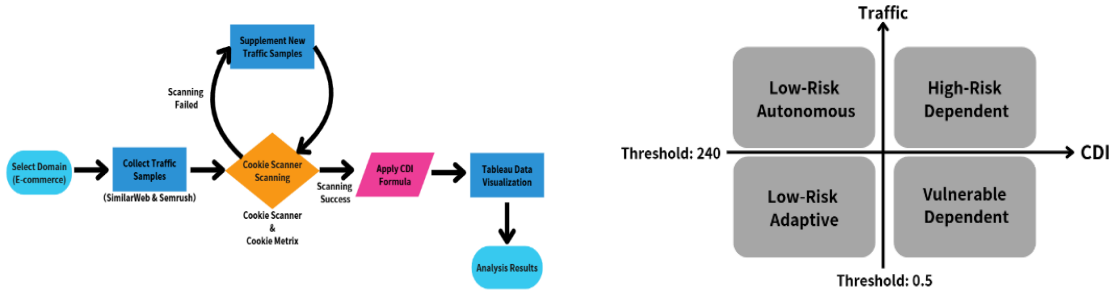


Fig. 1: Overall Research Methodology Flowchart Fig. 2: Quadrant Classification of CDI Risk Levels

2.4. Quadrant Risk Assessment Model

To visually analyze the relationship between CDI and website traffic, this study employs a quadrant analysis method to evaluate different website risk types. The x-axis represents CDI scores, with 0.5 as the threshold for high/low dependency, while the y-axis uses an average traffic threshold of 240M visits across the 100 websites. Websites are categorized into four risk levels, as shown in Figure 2:

- High-Risk Dependent (High Traffic & High CDI): Highly trafficked websites but are highly dependent on third-party cookies, facing severe tracking and revenue challenges post-cookie phase-out.
- Vulnerable Dependent (Low Traffic & High CDI): Low-traffic websites heavily reliant on third-party cookies. The loss of third-party cookies could severely impact these websites, making them the most vulnerable category in this quadrant analysis.
- Low-Risk Autonomous (High Traffic & Low CDI): High-traffic websites and low dependency on third-party cookies. This suggests reliance on first-party data, alternative tracking methods or strong brand influence, making them less affected by third-party cookie elimination.
- Low-Risk Adaptive (Low Traffic & Low CDI): Low-traffic websites and low dependency on third-party cookies, possibly independent of third-party cookies and least affected by their removal.

2.5. Tableau Quadrant Chart

To enhance the clarity of CDI and website traffic relationships, this study employs Tableau for data visualization and quadrant analysis (Figure 2). Tableau is a powerful tool capable of handling large datasets, allowing researchers to customize thresholds. This study constructs the CDI quadrant chart and classifies websites into four risk categories.

Since the collected traffic data ranges from 33.09M to 5450M visits, plotting on the original scale compresses low-traffic website variations, making quadrant classification unclear—especially for Low-Risk Adaptive and Vulnerable Dependent sites. This distorts data distribution readability. To address this issue, Prior research [13] suggests that a logarithmic scale prevents small values from being overly compressed when data spans multiple magnitudes, improving trend visibility and comparability. Therefore, the chart applied a logarithmic scale ($\log_{10}$) to the y-axis (traffic) to maintain analytical accuracy and readability.

3. Results & Discussion

This chapter compares CDI quadrant visualizations using original vs. logarithmic scales, analyzes the CDI-traffic relationship and evaluates the impact of third-party cookie deprecation. It also examines strategic responses, Privacy Sandbox implications, and market competition issues.

3.1. Visualizing CDI and Traffic: Original vs. Logarithmic Scale

Figure 3 displays website traffic and CDI on the original scale. This approach stretches the graph due to high-traffic websites (exceeding 1000M), clustering low-traffic sites and obscuring quadrant distinctions, particularly between Low-Risk Adaptive and Vulnerable Dependent categories

Figure 4 applies a logarithmic scale, evenly distributing data points and improving visibility of low-traffic CDI variations. Compared to the original scale, this enhances risk classification clarity and prevents high-traffic dominance from skewing data distribution.

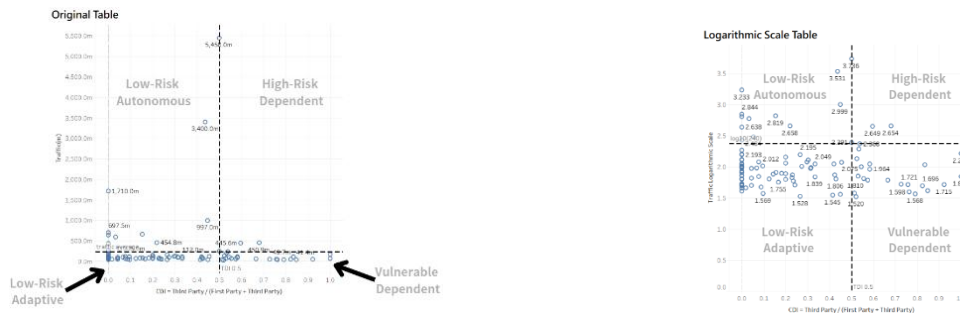


Fig. 3: CDI and Website Traffic (Original Scale) Fig. 4: CDI and Website Traffic (Logarithmic Scale)

3.2. Risk Classification and Case Analysis

This section presents case studies from different e-commerce websites representing each of the four risk quadrants. By analyzing their specific cookie reliance, business risks, and adaptation strategies, this study provides clearer insights into the challenges these platforms may face following third-party cookie deprecation, offering valuable references for e-commerce platforms and digital marketers.

- High-Risk Dependent (High Traffic & High CDI):

Rakuten.co.jp, a major Japanese e-commerce platform, relies heavily on cross-platform user tracking and personalized recommendations to boost engagement and conversion rates. When third-party cookies are deprecated, these types of websites will likely face reduced ad precision, tracking restrictions, and declining conversion rates, making it imperative for them to adapt to alternative tracking and targeting methods[22].

- Vulnerable Dependent (Low Traffic & High CDI):

Promodescuentos and Kabum represent lower-traffic sites heavily reliant on third-party cookies for behavioral tracking and personalized marketing. Due to their relatively lower traffic compared to major e-commerce platforms, such as Amazon, the phase-out of third-party cookies will likely diminish their ad effectiveness and user retention, driving their traffic further toward larger competitors. Among the four quadrants, this segment faces the most significant risks.

- Low-Risk Autonomous (High Traffic & Low CDI):

eBay, a globally recognized e-commerce platform, has relatively low dependency on third-party cookies by extensively using first-party shopping insights for targeted advertising and privacy compliance[23]. This deep integration of first-party data reduces eBay's reliance on third-party tracking technologies, positioning it as highly resilient to the challenges posed by the impending deprecation of third-party cookies, representing the most favorable scenario among the four quadrants.

- Low-Risk Adaptive (Low Traffic & Low CDI):

While these platforms have relatively lower online traffic compared to major competitors, they typically leverage first-party data or offline channels to maintain competitiveness. A representative example is Costco, which recently launched its Retail Media Network (RMN), capitalizing on the purchasing history and preferences of its extensive membership data to deliver highly targeted advertisements and promotions, thereby significantly reducing their dependency on third-party cookies[24].

3.3. Adaptation Strategies in the Cookieless Era

With the third-party cookies phase-out, businesses and digital advertisers must adapt marketing strategies. However, this does not signal the end of personalized marketing. Key approaches include:

- First-party data collection: Businesses will leverage CRM systems, membership programs, and direct user engagement to gather more accurate data, fostering stronger customer relationships and reducing dependence on third-party tracking [14].
- Contextual Advertising: AI-driven contextual ads target users based on website content rather than browsing history, ensuring privacy compliance and benefiting both enterprises and SMEs[15].
- Server-Side Tracking: This method sends user data directly to servers, reducing reliance on browser-based tracking and enhancing data security [16].
- SEO and Content Marketing: SEO has emerged as a critical digital marketing strategy by optimizing website content, implementing strategic keywords, and making technical adjustments to boost search engine rankings. This approach allows businesses to attract their target audience naturally without relying on third-party cookie tracking. Additionally, content marketing offers an effective alternative by delivering high-quality content that attracts users and fosters brand loyalty. [17]

3.4. Privacy Sandbox and Antitrust Controversy: Market Impact and Regulatory Concerns

Google's Privacy Sandbox aims to replace third-party cookies with privacy-preserving ad solutions (Topics, TURTLEDOVE, Dovekey), sparking competition concerns [18].

- Google's market dominance: With third-party cookies phased out, Google's monopolistic position may be further reinforced, leading advertisers to struggle to target users outside Google's ecosystem, while Google retains vast first-party data from Search, YouTube, etc, increasing its ad dominance .[19]
- Walled gardens & competition barriers: Google-controlled alternatives restrict independent ad tech firms' access to user data, reinforcing a closed ecosystem that limits competition and innovation. [20].
- Regulatory scrutiny: Agencies like EU Commission and UK's CMA are investigating Google's self-preferencing in ad auctions and potential anti-competitive practices in Chrome's ad framework. [20]

4. Conclusion & Future Work

The Cookie Dependency Index (CDI) provides a quantifiable measure of third-party cookie reliance, identifying high-risk websites and guiding adaptation strategies. CDI serves as a valuable tool for businesses and researchers in data dependency assessment and decision-making. Future research can enhance data accuracy and risk evaluation by:

- Refining traffic analysis or integrating internal data to reduce third-party tool errors.
- Exploring CDI variations across different industries and regulatory impacts (e.g., GDPR, CCPA).
- Improving risk classification through advanced visualization or machine learning.
- Investigating profitable low-traffic websites is worthwhile, as some such websites exhibit unexpectedly high profitability. This phenomenon is worth further examination, as it may be related to advanced data utilization, subscription-based business models, or personalized advertising strategies.

As digital advertising and privacy policies evolve, refining risk assessment frameworks is crucial. CDI offers a structured approach to evaluating third-party cookie dependency and supporting business decision-making. Future research should track emerging technologies and policy shifts to enhance market adaptability, while exploring alternative solutions to sustain a competitive edge in data-driven marketing

5. Acknowledgements

Angus Chen would like to express his gratitude to Professor Jiang-Yu Zheng for his guidance and encouragement, which enabled him to successfully complete this paper.

6. References

- [1] CAHN, Aaron, et al. An empirical study of web cookies. In: Proceedings of the 25th international conference on world wide web. 2016. p. 891-901.
- [2] KAUR, Ms Shagundeep. (2022). COOKIES: A CONCEPTUAL STUDY.

- [3] JONES, Meg Leta. The Character of Consent: The History of Cookies and the Future of Technology Policy. MIT Press, 2024.
- [4] DEGELING, Martin, et al. We value your privacy... now take some cookies: Measuring the GDPR's impact on web privacy. arXiv preprint arXiv:1808.05096, 2018.
- [5] WAGNER, Paul. Cookies: Privacy risks, attacks, and recommendations. Attacks, and Recommendations (December 8, 2020), 2020.
- [6] Google, "Prepare for Third-Party Cookie Deprecation," Google Ads Help.[Online].Available:<https://support.google.com/google-ads/answer/14762010?hl=en>.
- [7] Google, "The Privacy Sandbox: A Safer Web for Users and Businesses," Privacy Sandbox. [Online]. Available: <https://privacysandbox.com/open-web/#the-privacy-sandbox-timeline>.
- [8] RUZIVE, Benson; MASENGU, Reason; MUCHENJE, Chenjerai. Reduction of Third-Party cookies—its effects in digital Marketing Transformation. Business Transformation-Accelerators for Sustainable Growth, 2023, 245.
- [9] Cookie Scanner. (n.d.). Cookie Scanner: Free Cookie and Privacy Audit Check. <https://cookie-scanner.com/>
- [10] CookieMetrix. (n.d.). Free Online EU Cookie Law Analyzer. <https://www.cookiemetrix.com/>
- [11] TAY, Yi Hong, et al. Ensuring Privacy and Security on Banking Websites in Malaysia: A Cookies Scanner Solution. Journal of Informatics and Web Engineering, 2023, 2.2: 153-167.
- [12] ENGLEHARDT, Steven; NARAYANAN, Arvind. Online tracking: A 1-million-site measurement and analysis. In: Proceedings of the 2016 ACM SIGSAC conference on computer and communications security. 2016. p. 1388-1401.
- [13] FIELD, James A. Some advantages of the logarithmic scale in statistical diagrams. Journal of Political Economy, 1917, 25.8: 805-841.
- [14] LÖNBERG, Nora; ANDO MANTOVAARA, Ludvig; JOHANSSON, Pontus. The Death of the Third-Party Cookie: A qualitative study regarding new strategies and methods for collecting consumer data to obtain effective digital marketing. 2022.
- [15] ZHANG, Kaifu; KATONA, Zsolt. Contextual advertising. Marketing Science, 2012, 31.6: 980-994..
- [16] ELMÉR, Johanna; NILSSON, Johanna. A FUTURE WITHOUT THIRD-PARTY COOKIES A study of how Swedish small and medium-sized marketing agencies are affected by the loss of third-party cookies and how potential change strategies are communicated. 2023. Master's Thesis.
- [17] YALÇIN, Nursel; KÖSE, Utku. What is search engine optimization: SEO?. Procedia-Social and Behavioral Sciences, 2010, 9: 487-493.
- [18] MENDYS, Andreas; JENSEN, Jens. How will Google Privacy Sandbox impact e-CRM of Danish SMEs. A single case study with Freeway ApS MSC INTERNATIONAL MARKETING MASTER THESIS, 2021.
- [19] COOPER, Dylan A., et al. Privacy considerations for online advertising: A stakeholder's perspective to programmatic advertising. Journal of Consumer Marketing, 2023, 40.2: 235-247.
- [20] GERADIN, Damien; KATSIFIS, Dimitrios; KARANIKIOTI, Theano. Google as a de facto privacy regulator: Analyzing Chrome's removal of third-party cookies from an antitrust perspective. 2020.
- [21] MILLER, Klaus M.; SKIERA, Bernd. Economic consequences of online tracking restrictions: Evidence from cookies. *International journal of research in marketing*, 2024, 41.2: 241-264.
- [22] OKSANEN, Timo. Companies' maturity for the deprecation of third-party cookies. 2022.
- [23] eMarketer. (n.d.). *How eBay Ads Create a More Satisfying and Successful Shopping Experience—It's in the Data (Sponsored Content)*. Retrieved from <https://www.emarketer.com/content/how-ebay-ads-create-more-satisfying-successful-shopping-experience-it-s-data-sponsored-content>
- [24] Adgate, B. (2024, June 12). *Costco Announces It Will Launch a Retail Media Network*. Forbes. Retrieved from <https://www.forbes.com/sites/bradadgate/2024/06/12/costco-announces-will-launch-a-retail-media-network/>