Poster: Advances and Open Problems in Web Tracking

Yash Vekaria*, Yohan Beugin*, Shaoor Munir, Gunes Acar, Nataliia Bielova, Steven Englehardt †, Umar Iqbal, Alexandros Kapravelos, Pierre Laperdrix, Nick Nikiforakis, Jason Polakis, Franziska Roesner, Zubair Shafiq, Sebastian Zimmeck

University of California, Davis *University of Wisconsin–Madison* *Radboud University*
Inria Centre at Université Côte d'Azur † *Independent Researcher* *Washington University in St. Louis*
North Carolina State University *Centre National de la Recherche Scientifique* *Stony Brook University*
University of Illinois Chicago *University of Washington* *Wesleyan University*

***Abstract*—Web tracking is a pervasive and opaque practice that enables personalized advertising, retargeting, and conversion tracking. Over time, it has evolved into a sophisticated and invasive ecosystem, employing increasingly complex techniques to monitor and profile users across the web. The research community has a long track record of analyzing new web tracking techniques, designing and evaluating the effectiveness of countermeasures, and assessing compliance with privacy regulations. Despite a substantial body of work on web tracking, the literature remains fragmented across distinctly scoped studies, making it difficult to identify overarching trends, connect new but related techniques, and identify research gaps in the field. Today, web tracking is undergoing a transformation, driven by fundamental shifts in the advertising industry, the adoption of anti-tracking countermeasures by browsers, and the growing enforcement of emerging privacy regulations. This poster presents a recently conducted Systematization of Knowledge (SoK) that consolidates and synthesizes this wide-ranging research, offering a comprehensive overview of the technical mechanisms, countermeasures, and regulations that shape the modern and rapidly evolving web tracking landscape. Our work also highlights open challenges and research directions, aiming to serve as a unified reference and introductory material for researchers, practitioners, and policymakers alike.**

Preprint. Yash Vekaria*, Yohan Beugin*, Shaoor Munir, Gunes Acar, Nataliia Bielova, Steven Englehardt, Umar Iqbal, Alexandros Kapravelos, Pierre Laperdrix, Nick Nikiforakis, Jason Polakis, Franziska Roesner, Zubair Shafiq, and Sebastian Zimmeck. SoK: Advances and Open Problems in Web Tracking, 2025. arXiv:2506.14057

1. Introduction

Users access a variety of free content and services on the web largely funded through online advertising which is itself heavily dependent on monitoring users' online activities. Ever since the introduction of cookies on the web in the mid-1990s, web tracking has evolved into a significantly more prevalent and sophisticated practice [8]. Moreover, user tracking and profiling often involves collection of user's personal details (e.g., name, email, and location), device characteristics (e.g., device model and operating system), browsing history, and behavioral signals (e.g., time spent on a page and performed interactions). As a result, web tracking has become an active area in online privacy research.

Over the years, researchers have conducted numerous studies to examine the evolution of web tracking mechanisms, browser developments, and regulatory compliance. Most survey papers have either been too high-level [4] or too narrowly-focused on specific forms of web tracking such as caching and browser fingerprinting [2], [6]. Overall, despite a considerable body of work, major findings remain scattered across many disparate studies. Furthermore, as privacy defenses improve in browsers, trackers continually adapt with new evasion techniques [7], [9]. The result is an ever-shifting technical landscape of tracking techniques. Similarly, regulations have continued to reshape the ecosystem by often governing tracking practices and ensuring that browsers provide necessary protections to safeguard user privacy. Today, web tracking is undergoing a transformative change due to the introduction of privacy-enhancing protections in major web browsers, new advertising paradigms, and evolving regulatory frameworks [3], [5].

These shifts call for a comprehensive and systematic study of emerging trends in the evolving tracking landscape to identify crucial research gaps. Thus, the community can clearly benefit from a unified resource that consolidates and systematizes the state of knowledge, helping researchers to make meaningful contributions to the field and ensure a structured approach at addressing new privacy issues.

This poster presents the following main contributions from our SoK (see reference above):

- We systematically organize the extensive body of research on web tracking, providing a consolidated knowledge base of advances in the field, highlighting evolving trends, bridging emerging but related tracking mechanisms and identifying gaps in the field.

*. Equal contribution. Contact for this poster: ybeugin@cs.wisc.edu

- We provide an overview of major browser-based anti-tracking interventions and relevant regulations across the EU and the US to assess how they have altered the ecosystem over the years.
- We identify key open challenges and promising future directions in the domain of web tracking for the community to address in the coming years.

2. Methodology

Online tracking has a vast literature, comprising numerous research studies published over the last few decades. As a result, we first carry out a literature survey to identify all papers related to web tracking published in the last 20 years (2005 onward) at any of the seven top web security and privacy venues—IEEE S&P, USENIX Security, ACM CCS, NDSS, ACM IMC, PETS, and WWW. A total of 200+ research papers were identified. Each paper was assigned one or more topics related to web tracking based on their abstract. The assignment of topics was jointly performed by two researchers following Clarke and Braun's thematic analysis approach [1]. A total of 40 topical themes were identified[1], with the top 15 (by number of papers) being tracking measurements, tracking in browsers and mobile, browser fingerprinting, regulation compliance, profiling, third-party tracking, cookie consent, cookies, user studies in tracking, ad blocking, JavaScript-based tracking, side-channels, advertising and tracking defenses, and browser extension fingerprinting. We used our domain expertise to structure the SoK around these prominent themes. Our work is scoped to how the data is *collected* about users, not how that data might then be *used*.

3. Main Takeaways from our SoK

Decades after its introduction, web tracking still remains an archetypal cat-and-mouse game. Each incremental defense—whether a new browser policy or a regulatory ruling—quickly provokes an equally sophisticated evasion technique to track users. This adversarial dynamic shows that purely reactive approaches cannot deliver privacy guarantees for online users.

Regulations alone are insufficient; data protection statutes such as GDPR and CCPA have tightened accountability, yet enforcement lags the speed of technical changes in evolving tracking mechanisms. Moreover, trackers often find tolerated gray zones to bypass regulations. As a result, enforcement frequently stalls on jurisdictional or interpretative disputes. Regulators need to incorporate agile, evidence-driven auditing methods by collaborating with the measurement community to avoid any oversight and to ensure that regulations evolve competitively with the technical reality.

On the other hand, while browsers are powerful gatekeepers, they provide an unreliable line of defense. Default protections vary widely across browser vendors, experimental features sometimes ship years after the issues are identified, and commercial incentives often result in more permissive designs. Future research must therefore look beyond "*fix it in the browser*" remedies and explore complementary approaches that truly safeguard users' privacy. Thus, the current tracking landscape demands a default *privacy-first* solution where users can control their privacy as opposed to browsers or regulators. Our work highlights this by summarizing important findings in the evolution of web tracking and its prevention across the years and suggesting key future directions.

1. The thematic organization is available at https://osf.io/zsy7e/overview?view_only=a346b98bf89244b5b17d842f6fb3fb13

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Numbers 2041894, 2138138, 2047260, 2343611, and 2143363. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work was also supported in part by the Semiconductor Research Corporation (SRC) and DARPA, the Agence Nationale de la Recherche through the ANR-21-CE39-0019 FACADES and the ProjetIA-22-PECY-0002 iPoP projects, and ANR MRSEI TULIP (ANR-24-MRS0-0004), Inria DATA4US Exploratory Action project.

References

[1] Virginia Braun and Victoria Clarke. Successful Qualitative Research: A Practical Guide for Beginners. 2013. ISBN 978-1-84787-581-5

[2] Tomasz Bujlow, Valentín Carela-Español, Josep Solé-Pareta, and Pere Barlet-Ros. A Survey on Web Tracking: Mechanisms, Implications, and Defenses. *Proceedings of the IEEE*, 2017. https://doi.org/10.1109/JPROC.2016.2637878

[3] Anthony Chavez. Update on Plans for Privacy Sandbox Technologies, 2025. https://privacysandbox.google.com/blog/update-on-plans-for-privacy-sandbox-technologies

[4] Tatiana Ermakova, Benjamin Fabian, Benedict Bender, and Kerstin Klimek. Web Tracking - A Literature Review on the State of Research. *In Hawaii International Conference on System Sciences*. 2018. https://doi.org/10.24251/HICSS.2018.596

[5] James Hercher. The W3C Ad Privacy Group Taking The Little-Engine-That-Could Path To Success, 2022. https://www.adexchanger.com/ad-exchange-news/the-w3c-ad-privacy-group-taking-the-little-engine-that-could-path-to-success/

[6] Pierre Laperdrix, Nataliia Bielova, Benoit Baudry, and Gildas Avoine. Browser Fingerprinting: A Survey. *ACM Transactions on the Web*, 2020. https://doi.org/10.1145/3386040

[7] Arvind Narayanan. The Web Tracking Arms Race: Past, Present, and Future. *In Enigma*. 2018. https://www.usenix.org/conference/enigma2018/presentation/narayanan

[8] Tobias Urban, Yash Vekaria, Zubair Shafiq, Chris Böttger, and Barry Pollard. The 2024 Web Almanac: Third Parties. In The 2024 Web Almanac. 2024. https://almanac.httparchive.org/en/2024/third-parties

[9] Tim Vlummens, Aniketh Girish, Nipuna Weerasekara, Frederik Zuiderveen Borgesius, Gunes Acar, and Narseo Vallina Rodriguez. Bridges to Self: Silent Web-to-App Tracking on Mobile via Localhost. *35th USENIX Security Symposium (USENIX Security 26)*, 2026

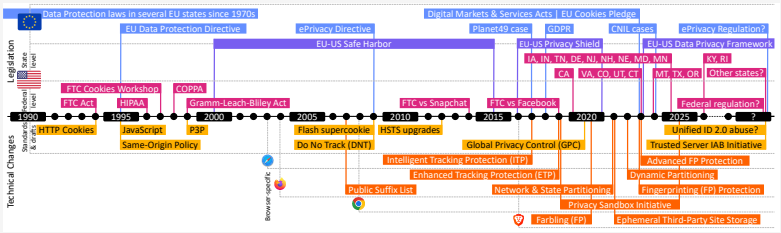
Advances and Open Problems in Web Tracking

Yohan Beugin - University of Wisconsin–Madison

MADS&P

Motivation

Web tracking is a pervasive and opaque practice

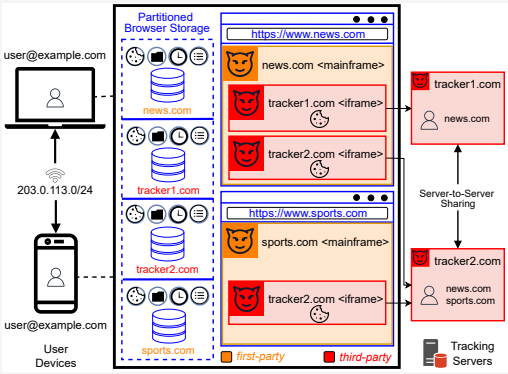


- Enables advertising, retargeting, conversion tracking, etc.
- Decades of research and regulations, but literature remains fragmented.
- Undergoing transformative changes with new protections and regulatory frameworks.

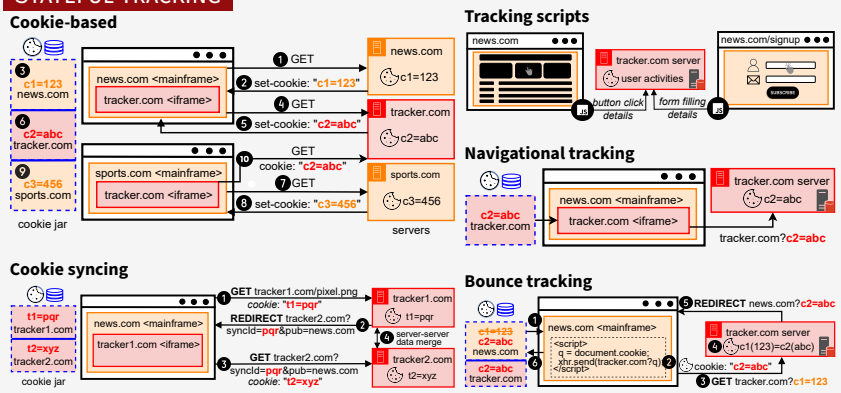
Need for a comprehensive and systematic study of emerging trends in the evolving tracking landscape to identify crucial research gaps

Threat Model

Browser's security model: context-origin boundaries



Stateful Tracking



Defenses: clearing and restricting cookies, blocking trackers, limiting lifetime of first-party storage, removing identifiers passed in URL parameters.

Shift to first-party cookies & partitioning: what alternative forms of user tracking are emerging? that could increase privacy risks, and how might these risk manifest?

First-party data reliance & identity graphs: how can we detect or infer opaque first-party data flows within the identity provider ecosystem and quantify its privacy risks?

Tracking tags: how are they configured differently across sites, how does it impact tracking?

Cross-device Tracking

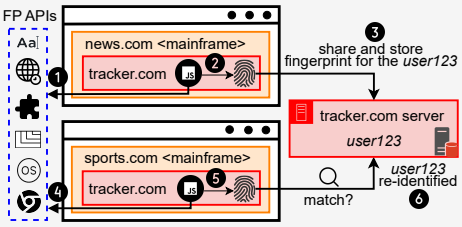
Deterministic or probabilistic

Defenses: account logins inherently limit protections, otherwise traditional defenses apply.

Theory to practice: how to characterize cross-device tracking in practice and defend against it?

Stateless Tracking

Fingerprinting (browser, side-channel, extension, behavioral, hardware)



Defenses: normalization, randomization, and anti-fingerprinting policies.

Paywalls: how are paywalls leveraged by publishers for building first-party profiles and perform targeted advertising within their own network?

Server to server: how does server-side tracking work, and how to detect and mitigate it effectively?

Fingerprinting: what is its real impact at scale, on vulnerable populations, and with other techniques?

Intent: how to distinguish between benign and tracking use cases of fingerprinting?

Hardware: How to prevent and detect low-level hardware-based fingerprinting?

Browsers: How can browsers defend against emerging side-channels?

Other Open Problems

Current focus on privacy policy and consent: what about other types of compliance (e.g., EU-US data transfer law)?

Transdisciplinary studies: how to reconcile web actors' incentives, responsibilities, and users' expectations?

Prevention: can adaptive measurement, monitoring, and disclosure methods be developed to stay ahead of tracking tactics?

Browsers: what new privacy-preserving advertising technologies can be built by learning from issues in prior proposals? How can we reliably and at scale automate the evaluation of their potential risks?

Other ecosystems: how do tracking mechanisms and protections diverge across web and app ecosystems?

Generative AI: how can we counter the privacy, security, and safety risks amplified from the use of generative AI by browsers, websites, and embedded third-parties for tracking, profiling, and personalization?

Takeaways

Archetypal cat-and-mouse game

- Browsers are powerful, but unreliable, gatekeepers.
- Regulations alone are not enough (slow enforcement vs. new ways).

Purely reactive approaches are insufficient

- Need for collaboration between regulators and measurement community (agile and evidence-driven auditing).
- Default privacy-first solutions for users to control their privacy.

Paper: SoK: Advances and Open Problems in Web Tracking

Authors: Yash Vekaria*, Yohan Beugin*, Shaoor Munir, Gunes Acar, Nataliia Bielova, Steven Englehardt, Umar Iqbal, Alexandros Kapravelos, Pierre Laperdrix, Nick Nikiforakis, Jason Polakis, Franziska Roesner, Zubair Shafiq, Sebastian Zimmeck (* = equal contribution)

https://yohan.beugin.org yohhaan yohan@beugin.org