



Your Customer's Web Browser Is Your Greatest Vulnerability

ONLINE DATA PRIVACY REPORT

OCTOBER 2022

LOKKER has identified over 5.1 million data privacy risks.

Millions of Hidden Pixels, Trackers, and Tags Are At The Heart of Web Privacy Lawsuits.

LOKKER's latest analysis of 170,000 websites around the world turned up over 5.1 million data privacy risks lurking beneath the surface of company websites. The major culprits are third-party JavaScript trackers, fingerprinters, data skimmers, and session replay scripts that, while some provide beneficial website features, they may also collect and share the visitors' information, often with unauthorized partners and without the website owner's knowledge.

LOKKER sees the recent burst of data privacy-related class-action lawsuits ([California "Invasion of Privacy,"](#) [Hospitals' Online Data Sharing,](#) [Pennsylvania Wiretapping/Session Recording](#)) and enforcement action against [Sephora](#) by the California Attorney General for violations of the California Consumer Protection Act (CCPA) as the beginning of an expansion of US regulatory actions, not unlike GDPR enforcement in Europe. Getting control of website data privacy has never been more urgent for companies.



And while Facebook and Oracle are in the headlines for their widespread data collection practices, LOKKER research has found that Twitter, Microsoft, SnapChat, Pinterest and TikTok are all collecting data across Education, Financial Services, and Healthcare sites in alarming numbers.

So what's the issue?

Companies are inadvertently sharing personal data with a host of third parties, and thus putting themselves and their customers at risk – risking millions of dollars in legal expenses as well as regulatory fines and penalties. Not to mention widespread erosion of consumer trust.

Without visibility into the website privacy risks brought by third-parties, companies are exposed to a growing number of class-action lawsuits claiming violations of the current and imminent data privacy laws in California, Utah, Virginia, Connecticut and Colorado, not to mention existing federal wiretapping laws.



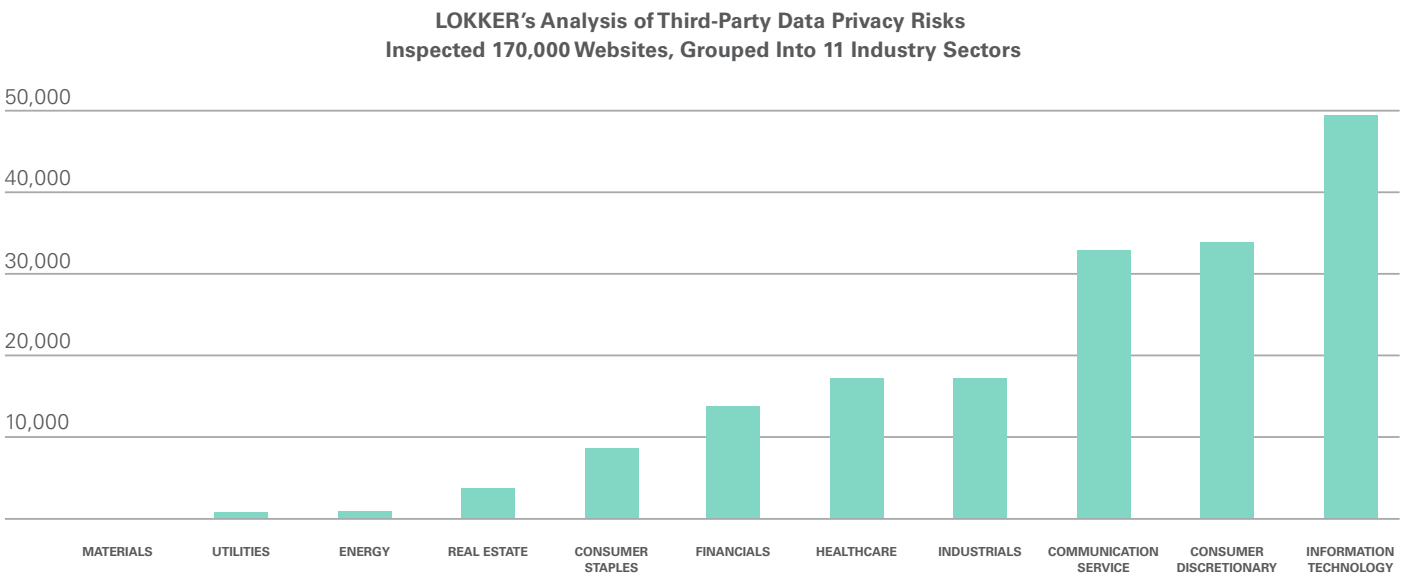
The Web Browser Is the New Endpoint to Defend

Taking a page from the cyber security playbook, savvy organizations are extending their data privacy efforts all the way to the ‘edge.’ With the ever-increasing reliance on third-party applications (e.g., ads, chatbots) and cloud services (eg., video delivery), websites introduce a host of security and data privacy threats. Thus, the customer’s web browser is now a hotbed of cyber risk, exposing visitors to malware, theft of their private information, and violations of privacy laws. But who’s protecting that endpoint?

According to [HTTP Archive’s latest “Annual State of the Web Report” \(Sep. 2022\)](#), 94% of sites use at least one third party and, on average, the top 1,000 websites use 53 third-party scripts including ads, analytics, CDNs, content providers, and social media features. With all this activity being introduced into millions of browser sessions, it has become mayhem for unauthorized data collection, theft and exploitation.

This report will quantify online privacy risks, describe the indicators of data insecurity and threats posed to organizations and customers, and provide guidance on what organizations can do to better protect themselves and their customers.

[LOKKER’s first edition of our online data privacy report \(reported in March, 2022\)](#) analyzed over 90,000 sites. For the second edition, LOKKER extended the research and analyzed privacy risk across 170,000 websites. Sources included Alexa, Cisco Umbrella and DomCop, and categorized each site into one of 11 industry sectors.



The 9 Web Privacy Risks You Need to Get Under Control

LOKKER has identified nine categories of third-party scripts that could wreak havoc with companies web operations and compromise customers' privacy. LOKKER's research uniquely quantifies 'online privacy risk' by classifying third-party scripts into these categories.

Web Privacy Risk Categories	Description
Malware	Malicious script that could deliver ransomware or keystroke logger that captures log-in credentials
Data skimming of PII and PHI	Possible Magecart attack or unauthorized data sharing that leads to GDPR, CCPA or privacy policy violations
Trackers	JavaScript that includes some 'event' or 'action' script that collects information from the host site and transmits to a third-party. This is what often leads to unauthorized data sharing.
Cookies being set in each session	Third party cookies may not be authorized or violate privacy policy
Fingerprinting Scripts	Profiling of anonymous users that is used to link with personal identifiers by data brokers
Foreign domains making requests	Potential GDPR violation or risks by known nation-state actors
Session replay scripts recording activity	PII may be captured incidentally
Young domains serving JavaScript	Malicious scripts are often served by recently established domains
Bad SSL Certification	Domains serving scripts may not be secure and could indicate bad actors

Often, companies that employ these scripts:

- Have no visibility to the extent of the data collection nor telemetry (sending data offsite)
- Lose track of which scripts are still operating, beyond the term of a contract with that provider
- Have not configured features correctly and may be inadvertently collecting and sharing PII
- Cannot track the data that is shared with their legitimate partners that may also be shared with unauthorized 4th, 5th or Nth parties



93% of Online Trackers Are From Google, Facebook, and Microsoft

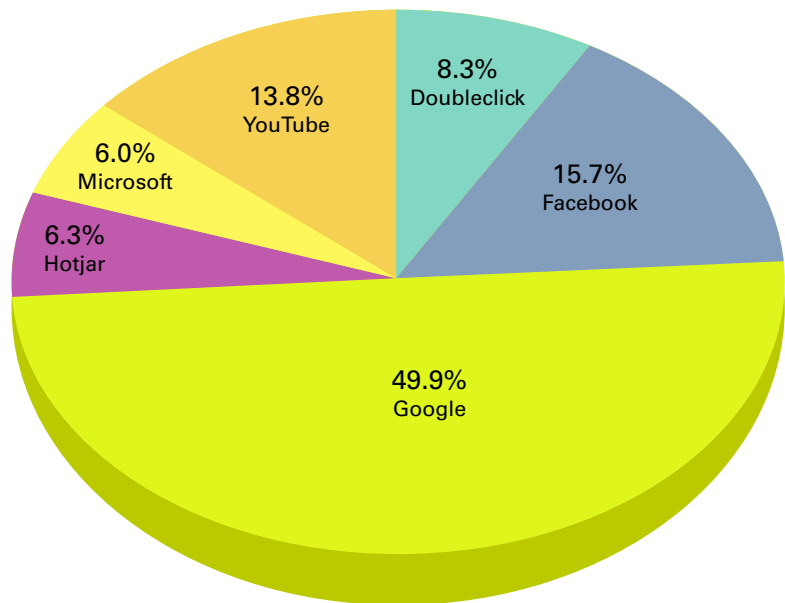
LOKKER's analysis of the data privacy elements outlined above, classified the vast majority as "trackers," or JavaScript that is collecting information on the host site and sending it to a third party (JavaScript events or 'actions' that include telemetry).



A recent **class action lawsuit alleges Facebook** used a web tracker to collect information from hospital website activities. According to the recent findings by The Markup, "A tracking tool installed on many hospitals' websites has been collecting patients' sensitive health information—including details about their medical conditions, prescriptions, and doctor's appointments—and sending it to Facebook." (The full Markup article can be found [here](#).)

It is no surprise that the most common third-party trackers, globally, were from the 3 online data giants, **Google, Facebook, and Microsoft**. This also includes sub-brands like Google's Doubleclick ad network and Microsoft's LinkedIn platform.

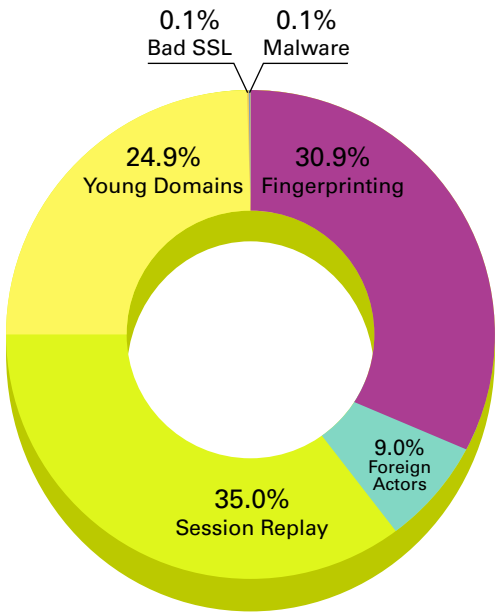
Top Online Trackers by Company



Rather unexpected, was the prominence of session replay scripts. Most common was from Hotjar, seen on almost 40,000 sites. HotJar is one of several types of "session recording" tools used by websites to record site visitor behavior to help site owners improve their user experience. However, if the tool is not configured properly, it can also be recording sensitive information. These session recorders are at the heart of the **recent data privacy lawsuit in Pennsylvania**, claiming that several companies have violated wiretapping laws.

In addition to the risks that the JavaScript ‘trackers’ present to companies, LOKKER’s analysis includes several other classifications of online data privacy threats, as follows:

Of the 170,000 Websites Scanned, The Following Data Privacy Threats Were Present



38,000+ Fingerprinting Scripts Lead to Widespread Consumer Profiling

As illustrated in the chart to the left, the proliferation of browser fingerprinting is significant and indicates that companies are becoming more sophisticated in creating alternatives to ‘cookie data.’ Fingerprinting scripts are able to capture details about the website visitor including location, IP address, type of device, fonts installed, and other specifications of their computer and browser. While this information alone is not ‘personally identifiable information,’ it is used by savvy data brokers to create profiles (‘fingerprints’) that are continually enriched until the visitor is able to be identified.

Over 11,000 Scripts Originated in Russia, Belarus, China and Iran

While the various types of third-party JavaScript requests are often served by CDNs for optimized delivery, LOKKER is also able to identify the location of domains serving these scripts directly. As many cyber attacks (ransomware, distributed denial-of-service) are often initiated by known nation-state actors, LOKKER has identified over 11,000 scripts originating from Russia, Belarus, China and Iran, the vast majority of which are from Russia.





Top 6 Social Media Networks Are Harvesting Data from Education, Financial Services, And Healthcare Sites

As mentioned in the introduction of this report, recent class action lawsuits filed in California alleged several healthcare organizations were sharing data with Facebook (Meta). LOKKER wanted to understand how widespread the issues were and uncovered that the “Meta Pixel” is not the only social media tracker deployed across hospitals, financial services, and educational sites.







Across the homepages of the **Fortune 1000** websites, Facebook trackers were identified on 46% of sites, Microsoft on 31%, Twitter on 21%, and Pinterest on 11%.

LOKKER’s analysis of the **Healthcare** industry (in the US) found the following: Facebook on 40%, Microsoft on 13%, Twitter on 8% and Pinterest on 6% of over 5,000 hospital and healthcare services (more details below).

Of note, TikTok trackers were discovered on 5% of the sites analyzed. When it comes to **Financial Services** sites in the US, LOKKER discovered that Facebook is on 36% of sites, Microsoft on 19%, Twitter on 10%.

LOKKER’s analysis of **Education** sites in the US (over 6,000 domains inspected) found that Facebook trackers are on 42% of sites, Microsoft on 15%, Twitter on 10%, and both SnapChat and TikTok on 5% of sites.

Presence of Social Media Trackers Across Sites in Key Industry Segments

SECTOR						
Fortune 1000	46%	31%	21%	11%	6%	7%
Healthcare	40%	13%	8%	6%	5%	3%
Financial Services	36%	19%	10%	2%	2%	1%
Education	42%	15%	10%	2%	5%	5%

The following provides details of these findings, based on the size of the organizations that were analyzed: Analyzing over 5,000 **healthcare and hospital** websites in the US, we found...

	Total	f	🎵	in	🐦	👤	📌
Health/Hospitals 10001+	256	45.28%	3.40%	30.19%	16.23%	3.40%	7.17%
Health/Hospitals 5001-10000	154	48.70%	3.90%	20.13%	15.58%	5.84%	7.14%
Health/Hospitals 1001-5000	486	51.03%	2.88%	23.46%	9.88%	1.65%	3.70%
Health/Hospitals 501-1000	266	46.99%	3.38%	24.44%	10.15%	2.26%	4.14%
Health/Hospitals 201-500	372	46.51%	3.23%	16.67%	9.14%	3.49%	4.84%
Health/Hospitals 51-200	868	45.51%	6.34%	18.55%	8.76%	3.46%	8.06%
Health/Hospitals 11-50	1558	42.75%	6.16%	11.17%	6.74%	2.44%	7.77%







Analysis of over 3,000 **Financial Services** sites identified the following...

	Total	f	🎵	in	🐦	👤	📌
Financial Services 10001+	53	64.15%	3.77%	43.40%	43.40%	9.43%	22.64%
Financial Services 5001-10000	43	65.12%	4.65%	46.51%	23.26%	6.98%	11.63%
Financial Services 1001-5000	251	51.39%	3.59%	38.25%	16.33%	1.59%	3.19%
Financial Services 501-1000	254	56.69%	2.36%	28.74%	10.63%	1.18%	1.97%
Financial Services 201-500	427	52.69%	2.11%	27.17%	10.54%	2.11%	1.87%
Financial Services 51-200	779	44.67%	2.57%	22.85%	9.64%	1.41%	1.80%



Some sites have as many as 232 cookies collecting visitor information.

LOKKER's inspection of over 6,000 education-related websites...








	Total						
Education 10001+	117	38.46%	11.11%	17.85%	20.51%	10.26%	2.56%
Education 5001-10000	142	59.15%	17.61%	31.69%	17.61%	16.20%	1.41%
Education 1001-5000	816	66.18%	14.09%	30.51%	16.18%	18.14%	1.47%
Education 501-1000	562	60.68%	9.61%	21.17%	12.10%	11.21%	1.07%
Education 201-500	600	53.67%	4.50%	16.33%	12.67%	7.67%	1.67%
Education 51-200	936	44.12%	2.56%	16.45%	8.23%	1.38%	2.35%
Education 11-50	1511	31.63%	1.79%	11.85%	9.53%	0.73%	1.59%
Education 1-10	1543	24.43%	0.97%	5.77%	7.00%	0.71%	1.75%

Oracle is Fueling the Controversial Data Broker Ecosystem.

As the social media giants continue to collect and share web visitor data, this past August, Oracle was also named in a class action lawsuit filed in California, alleging that the company "has amassed detailed dossiers on some five billion people, accusing the company and its adtech and advertising subsidiaries of violating the privacy of the majority of the people on Earth." (Full article [here](#)).

LOKKER research uncovered that the Oracle "Bluekai" tracker (one of the elements under scrutiny in this lawsuit) and its "AddThis" cookie are widely distributed on sites throughout the world.

Presence of Oracle Tracker and Cookie Discovered on LOKKER Inspected Sites

SECTOR							
Fortune 1000	46%	31%	21%	11%	6%	6%	7%
Healthcare	40%	13%	8%	6%	5%	5%	8%
Financial Services	36%	19%	10%	2%	2%	2%	7%
Education	42%	15%	10%	2%	5%	5%	7%

Once again, this represents that hidden third-party JavaScript may be performing a needed function on the host site, and yet collecting and sharing data with third parties in an unauthorized manner.

On Average, Sites Deploy 26 Cookies, Some Sites as Many as 322

LOKKER also analyzed the use of cookies across the 170,000 inspected websites. Cookies are an essential part of the data privacy conversation because cookies collect and store information (in the browser software), even after a visitor has left a site. The risk to companies, however, is three-fold:

1. If a company is required to obtain consent from site visitors to issue cookies, is the consent tool listing all of the actual cookies?
2. If personal data is being stored in the cookie, is it being accessed by unauthorized third parties?
3. If a visitor requests that data is not collected or stored, is that customer data removed from the company's records (or suppressed)?

Elegantly defined by Mozilla, cookies are mainly used for three purposes:

Session management

- Logins, shopping carts, game scores, or anything else the server should remember

Personalization

- User preferences, themes, and other settings

Tracking

- Recording and analyzing user behavior



As you will see in the chart below, cookies also provide **key information** that enable online ad targeting and tracking for companies like Google, Microsoft, Oracle, Casale Media, Bidswitch, etc. While some **sites had as many as 322 cookies**, we found, on average, that Fortune 1000 company homepages employ 26 cookies.

The 10 Most Common Cookies Are From These Domains:

Microsoft	.linkedin.com
Google	.doubleclick.net
AddThis (Oracle)	.addthis.com
Google	.youtube.com
Index Exchange	.casalemedia.com
Microsoft	.c.clarity.ms
Bidswitch	.bidswitch.net
AppNexus	.adnxs.com
Rubicon Project	.rubiconproject.com
Microsoft	.c.bing.com

The challenge for organizations in regards to cookie management and data privacy compliance, is to ensure that they:

- A. Have visibility to all of the cookies being placed into their customers' browser session
- B. That only the data they wish to collect and share is being used by authorized parties and in compliance with their stated privacy policy.



Prepare for 2023 Privacy Regulations

Given the recent class action lawsuits, the growing consumer concern about personal data privacy, and the imminent data privacy laws going into effect in 2023 in California, Virginia, Utah, Colorado, and Connecticut, there is increasing pressure on organizations to take control of data privacy risks on their websites.

For reference, new data privacy laws going into effect in 2023:

California	California Consumer Privacy Rights Act (CPRA) Jan 1, 2023
Virginia	Virginia Consumer Data Protection Act (VCDPA) Jan 1, 2023
Colorado	Colorado Privacy Act July 1, 2023
Utah	Utah Consumer Privacy Act December, 2023
Connecticut	Connecticut Data Privacy Act July 1, 2023

LOKKER recognizes that organizations have an immense responsibility to protect their customers and their company. In addition to constantly fending off threats from malicious actors, security, marketing, and privacy executives have the additional challenge of complying with data privacy laws.

To help organizations prepare for and mitigate data privacy risks, there is a strong ecosystem of attorneys, technology advisors, and software providers that can support their efforts. The [International Association of Privacy Professionals](#) is a great resource (IAPP.org). The team at The Markup also provides timely, valuable research and insights. Helpful information is also available at [The Electronic Frontier Foundation](#) and [The Center for Humane Technology](#).

LOKKER provides this online data privacy report twice each year to help organizations identify emerging data privacy threats and support their efforts to make web experiences safer for their customers. In addition to the research in the report, LOKKER provides its [Privacy Edge™ software](#) to organizations to continuously monitor, alert and block threats in the browser.

For more information on the findings in this report, or to learn more about LOKKER’s data privacy solutions, contact:

Jeremy Barnett
Chief Commercial Officer
jeremy@lokker.com
LOKKER.com