2020

The Internet and Web Tracking

Tim Zabawa
Grand Valley State University

Tim Zabawa

12/17/20

Capstone Project

Cover Page:

1. Introduction:

Can you remember the last time you didn't visit at least one website throughout your day? For most people, the common response might be "I cannot". Surfing the web has become such a mainstay in our day to day lives that a lot of us have a hard time imagining a world without it. What seems like an endless trove of data is right at our fingertips. On the surface, using the Internet seems to be a one-sided exchange of information. We, as users, request data from companies and use it as we deem fit. Most of us don't disprove of the transaction because we are on the receiving end of it. Unfortunately, what seems like a trade to our advantage actually requires us to give up quite a bit. What is given up for being able to utilize the web is information about oneself. This includes highly sensitive and personal data. The internet has come a long way since its infancy. As time has gone on, it has grown to be quite the complex application. Rooted in its capabilities are ways for advertisers to harvest and exploit information about consumers. This process is called web tracking.

Over time, the retrieval and consumption of user data has ballooned into a gigantic industry. It has spread its tentacles into almost every business sector. Companies use the information for a wide range of business decisions. The following study aims to break down several of the technologies used to track the average web client online, what they can do to defend themselves against such threats, and why their information is so sought after in the first place. Additionally, real world tests determining the effectiveness of several defensive measures will be provided. Because of the vastness of the topic, the study will only focus on the advertising aspect of web tracking. Furthermore, it does not intend to push users away from the Internet. It is to assist them in being more informed about how their data is serviced and the ways they can protect themselves.

2. How we are tracked online:

When a user surfs the web, they typically go through the following steps: first, the consumer uses a search engine such as Google or Microsoft Bing to view search results corresponding with the keywords typed into a search bar. Second, the user scans the viewable responses and eventually clicks on a website link. Third, the client is taken to the appropriate web page and utilizes the information for a given purpose. Fourth and final, the user then generally decides to either repeat steps one through three or discontinue using the Internet. Web

tracking can take place in more than one of the steps listed above. However, most of the data gathered on consumers will occur during step three. At the very moment a client first views a website, information on them is already being collected. The following paragraphs are going to break down three technologies used to track users online. They are cookies, browser fingerprinting, and web beacons.

2.1 Cookies:

One of the most common technologies used for web tracking is called a cookie. A cookie is a collective set of data points that can be utilized as a single unit or pulled apart and used as separate pieces of information. They are generated by the author of a website but managed by the user's web browser. The main purpose for a cookie is to enhance the experience a consumer has when he or she interacts with a website. Unfortunately, it is also commonly used for web tracking. Cookies come in three main types: session cookies, persistent cookies, and super cookies. Session cookies are automatically deleted when the user closes the browser [1]. The common use case for a session cookie would be to keep track of shopping cart items on an ecommerce website. Persistent cookies come in two categories: first-party and third-party. Both contain an expiration date. Because of this, they can last however long the owner dictates. The timeframe can even exceed a browser being closed and reopened at a later date [1]. The common use case for a persistent cookie would be to prevent a consumer from having to log back into their respective web accounts after they exit the browser. The difference between the two categories lies with whether the cookie originated from the current website or not. A first-party persistent cookie was created by the author of the page in view while a third-party persistent cookie was not. Super cookies not only act similar to the third-party persistent cookie but also extend access to a higher number of resources. They also are used for malicious purposes like changing user data or forging a login [1]. The third-party persistent cookie is synonymous with the "tracking cookie". When it comes to web tracking, users will more often than not be monitored by a third-party persistent cookie.

Cookies are created when a web page is first loaded inside a browser and displayed to the user. Along with the interface data, such as HTML markup, and CSS stylesheets, the response includes a cookie. The browser will then store the cookie inside a permanent storage location. This allows cookies to sit on device for potentially a very long time. As the user surfs the web,

data can be added to the cookie. This includes but is not limited to the history of viewed websites, how long they were viewed for, search queries, purchases, device information, when and where previous advertisements were seen, and location data [1]. Cookies are made up of numerous components. But the following ones are required: name, value, and attributes [1]. The "name" is used to identify which cookie is which and what it is used for. The "value" is used to uniquely identify the current consumer's browser. Lastly, "attributes" add additional pieces of information to the cookie. The expiration date of the cookie, whether the cookie is accessible by resources other than the originating website, and whether the cookie can be transferred over insecure networks are examples of cookie attributes [1].

Now that we know what a cookie is and how it makes its way onto a user's browser, in what ways can advertisers get their hands on the data? To assist in explanation, the following contrived example is provided: The owner of an ecommerce website has signed up for an advertising platform like Google. When a consumer visits the webpage, their browser picks up a Google ad's tracking cookie. That cookie contains a value property identifying the user's browser. When the consumer visits other websites that use the same advertising platform, the cookie is identified via the name property and sends it to Google. They now have access to all of the information contained within the cookie. From there, the value property of the cookie is traced back to the previous shopping site and an associative advertisement is displayed within the current website. Over time, tracking cookies can retrieve a lot of data about how a consumer utilizes the Internet. Both the author of the cookie as well as other websites use this information to create "profiles" on users. To make matters worse, this data is often collected without the consumers consent or knowledge. In the United Kingdom and European Union, websites are required by law to notify users if they use tracking cookies. In the United States and other countries, disclosure of tracking cookies is not required [1].

2.2 Browser Fingerprinting:

As users surf the web, websites are loaded in a browser. This program is curtailed to the operating system of the underlying computer. Examples include Google Chrome, Apple Safari, and Mozilla Firefox. These products offer a ton of settings, tools, and plugins to create a "personal" online experience. Many users may be surprised to know that the active features

associated with a browser is another avenue in which they can be tracked. This method of web tracking is called “browser fingerprinting”.

A browser fingerprint can be thought of as a signature made up of all of the data associated with a consumer’s browser. Because there are a small statistical chance two given users’ browsers share the exact same fingerprint, the unique signature can be exploited for advertising purposes. Panopticlick, a browser researching company, found that only 1 in 286,777 other browsers share the same fingerprint [2]. Stats such as this support why browser fingerprinting has become a very successful web tracking mechanism. Advertisers collect large sets of information on browsers. They can match those respective signatures in order to form both bulk and scoped user profiles [2]. Examples of what can be used to construct a browser fingerprint includes cookies, canvas fingerprinting, IP addresses, and JavaScript methods.

Cookies have already been aforementioned in the study. Canvas fingerprinting is a new method of obtaining browser data. Websites are written in HTML markup. It can be thought of as computer instructions to construct a web page that consumers can view and interact with. Inside that markup is an instruction that generates a “styled” piece of text. Along with the content are commands to make a request to the author of the website. When this request is made, browser information such as the type and version, operating system, active plugins, time zone, language, and screen resolution are shared with the recipient [2]. In contrast to how cookies work, canvas fingerprinting doesn’t include loading anything onto the computer. Therefore, it can’t just be deleted. IP addresses are used for computers to locate one another on the Internet. What users type into a search bar, such gvsu.edu or facebook.com, are transformed into a sequence of numbers and characters that a computer can understand. An example of an IP address would be 127.0.0.1. Computers use this address to connect the consumer with the correct website. Because IP addresses are unique to some degree, advertisers can potentially use the information to track users online. JavaScript is a programming language that runs in a browser. Depending on settings, a program written in JavaScript, called a script, can access a lot of data about the consumer’s internet activity. Browser fingerprinting is another method in which user information can be consumed. Advertisers use the essential data to construct profiles.

2.3 Web Beacons:

Cookies and browser fingerprinting have been identified as common web tracking methods. Consequently, numerous protective measures have been created over the years to mitigate their nefarious abilities. Advertisers have had to continuously come up with more innovative ways to consume user data. One alternative tool that has more recently gained traction is called a web beacon. Web beacons can also be referred to as web bugs or pixel tags [3]. They are small pieces of content that embed themselves inside a web page. From the consumer's perspective, they are invisible. Web beacons come in slight variations of a web page image. The most common form is the 1x1 GIF [3].

The popularity of web beacons has risen because of their effective ability to gather information on users. Unlike other web tracking schemes, web beacons gather data in a more efficient manner. The minute size of the image itself allows it to be downloaded extremely fast. This assists in not drawing the consumer's attention as he or she surfs the Internet. A slow loading website can often draw a red flag. Additionally, web beacons cannot easily be deleted. Because they are ingrained inside the raw markup of the web page, removing the image might cause negative consequences such as rendering issues. When the browser loads the web page and the content contains a web beacon, the browser makes a subsequent request to download the image. Similar to canvas fingerprinting, attached to this request is information on the underlying user and the user's browser. Web beacons can expose a consumer's IP addresses, the date and time, as well as reference any previous existing cookies that belong to the same website author [3]. Almost all information provided by web beacons is "non-identifiable". However, as more data is collected and analyzed across different web tracking avenues, advertisers gain a clearer picture of a what a user's online interests are.

3. Defenses Against Web Tracking:

It can be overwhelming for a user to realize that advertisers deploy such technological measures to track their internet actions and habits. But not all hope is lost. Over time, many organizations have tried to help individuals feel more at ease when surfing the web. This endeavor has produced numerous tools a user can deploy to use the Internet in a more private manner. The following paragraphs provide detail on several of the defensive measures one can enact against cookies, browser fingerprinting, and web beacons.

3.1 Cookies:

How can users defend themselves against the web tracking abilities of cookies? The first and likely the most straightforward action a consumer can take is removing cookies from the browser. All browsers offer the ability to delete cookies. A quick online search will provide numerous examples of step-by-step instructions as to how to remove undesired browser cookies. Consumers should note that browsers don't distinguish between the types of cookies upon deletion [2]. When cookies are cleared from a browser, all are destroyed.

The second defensive measure a user can activate is a "Do Not Track" browser setting. This action will prompt visited websites not to collect tracking information about a consumer. There is a caveat, however. The "Do Not Track" browser setting does not actually enforce web tracking guards nor is there lawful authority observation [2]. Websites can honor the privacy request if they would like but are not required. The browser setting is still a legitimate layer of protection, nevertheless.

The third method that can be implemented is to activate "incognito mode" or "private browsing" within a given browser session. Most, if not all of the most popular browsers offer this feature. Surfing the web in this mode restricts the browser from saving certain user data. This includes browsing history, cookies, site data, and form submission related information. Several browsers even enforce third-party cookie protections.

Finally, browser plugins can be activated to provide protection against tracking cookies. Privacy Badger is a well-known browser plugin for the Google Chrome and Mozilla Firefox browsers. It automatically blocks advertisers that use tracking cookies to load additional content in a browser. This is done by keeping track of third-party resources that embed images, scripts, and advertising into web pages [2]. The extension doesn't use a blacklist of known tracking web sites. Instead, Privacy Badger observes the behavior of third-party resources on web pages and blocks if they collect unique information about the consumer [2]. Disconnect is another example of a browser plugin [2]. It automatically detects when a browser requests a resource other than those belonging to the site being visited. After analyzing which assets are actually necessary for the web page, the extension will either block or allow the request. Users can alternatively choose to accept the requests on both a group and individual basis [2].

Defensive measures don't come without side effects. For example, browsers don't make distinctions between different types of cookies. When deleted, the "good" ones and the "bad"

ones are destroyed. This might affect previous account states held by users. Additionally, there have been many instances where browser plugins have introduced exploitable security vulnerabilities. The extensions may also increase the uniqueness of the underlying browser fingerprint as well as decrease downloading speeds. A consumer must accept a balance between utilizing some of these tools and an adequate internet experience.

3.2 Browser Fingerprinting:

How can users defend themselves against the web tracking abilities of browser fingerprinting? The goal of the following strategies is to "blend" the signature of one browser with the browsers of other online consumers. This will decrease the likelihood that a given profile of online data can be associated with a given user.

The first method that can be implemented is to activate incognito mode within a given browser session. This mode sets certain browsers' settings to a preconfigured set of values [2]. Since key aspects of a browser's fingerprint are made up from this information, it will be more likely that the generated signature will match other browsers.

Secondly, protective browser plugins can be installed. Similar to tracking cookies, extensions such as Privacy Badger and Disconnect monitor and disable browser signature data from reaching an undesired source [2].

Third, scripting languages such as JavaScript can be disabled. This will greatly reduce the detectability of information such as active plugins and fonts. Certain cookies will also be unable to get on the underlying browser [2]. Be aware that turning off JavaScript might prevent a website from working properly.

Fourth, anti-malware can add another layer of protection. Even though most anti-malware tools are not created to prevent browser fingerprinting, a lot of them will provide a defensive layer as a side effect. Common anti-malware features include blocking ads, toolbars, and spyware software that might be running in the background of a system [2]. These same perpetrators add to the uniqueness of a browser signature.

Finally, using a "virtual private network" (VPN) or the "Tor Browser" can help mask some browser fingerprinting data from advertisers. A VPN is like a middleman between two web sources trying to exchange information. Instead of connecting directly to a website, the user's browser will connect to the VPN's computer first. The machine will subsequently exchange the

requests and responses on behalf of the two web sources [2]. When a VPN is utilized, certain browser signature data of the original consumer is kept hidden. This type of guard can be taken a step further. The Tor Browser is an interface for accessing the Tor network, which is an alternative internet created for privacy and request anonymity [11]. Due to the fact that Tor uses certain default settings, it is harder for advertisers to identify unique browser fingerprints [11]. Additionally, the Tor Browser aggressively blocks JavaScript code on websites [11]. The Tor network is slower compared to the mainstream internet. It is also imperative that only the Tor Browser is used to access the Tor network and not browsers such as Google Chrome and Mozilla Firefox [11]. The latter will cause negative repercussions.

3.3 Web Beacons:

How can users defend themselves against the web tracking abilities of web beacons? One of the best mitigations comes indirectly from none other than Google. Not so long ago, the company started to implement security protections around caching images. Similar to anti-virus software, the primary goal for Google was to prevent security vulnerabilities. But the resolution also hindered web beacons. When an interface is identified as containing a picture, it is first downloaded, cached, and stored on a Google server [3]. When the consumer opens the contents of the web page, instead of downloading the image from the original source, it will be downloaded from Google [3]. This significantly impairs the usefulness of a web beacon because all of the collected data will be about Google and not the targeted user. As with tracking cookies and browser fingerprinting, browser plugins can add another layer of defense against web beacons. Ghostery is an extension that allows consumers to view and manage the discovered trackers on a given website [4]. This includes web beacons. The browser plugin provides the option of disabling a tracker on a single website or across all websites in which it is identified [4]. Unfortunately, web beacons are designed to be invisible and can't always be identified by privacy extensions.

4. Technological examples:

The examples provided are centered around https://www.dickssportinggoods.com. Dick's Sporting Goods, Inc. is an American sporting goods retail company, based in Coraopolis, Pennsylvania. As of 2018, it has approximately 850 stores and 30,000 employees [5]. Their

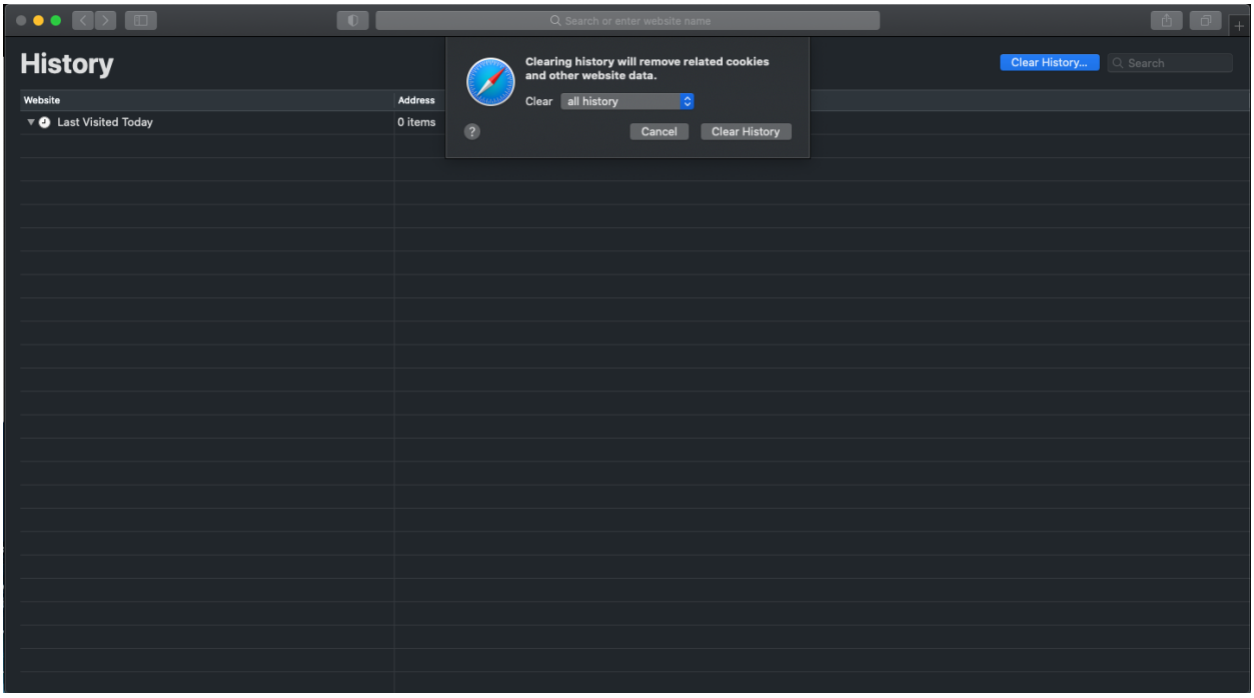
website ranks 2,296th out all global websites in terms of internet traffic and engagement. In the U.S., it is ranked 520th [6]. Based on informal testing, other websites with similar characteristics behave similarly with regards to their use of tracking cookies.

Example 1:

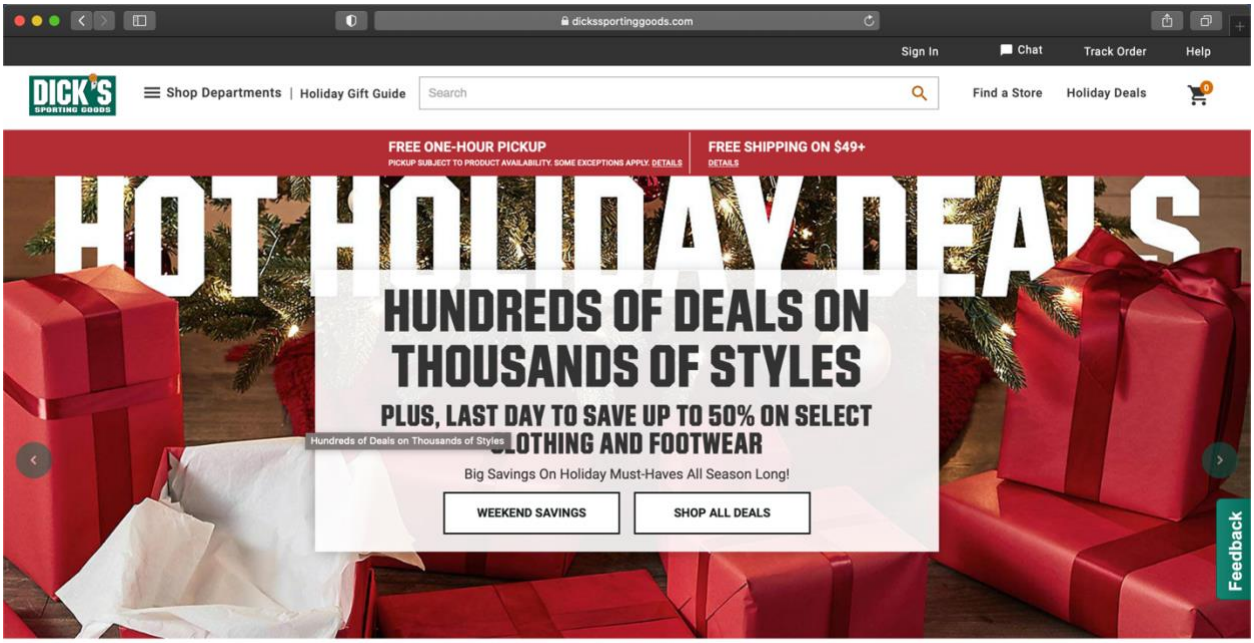
I carried out a test to determine how the Ghostery browser plugin protects users against the use of tracking cookies on an ecommerce website.

Test 1.1 reproduction steps (plugin **is not** installed):

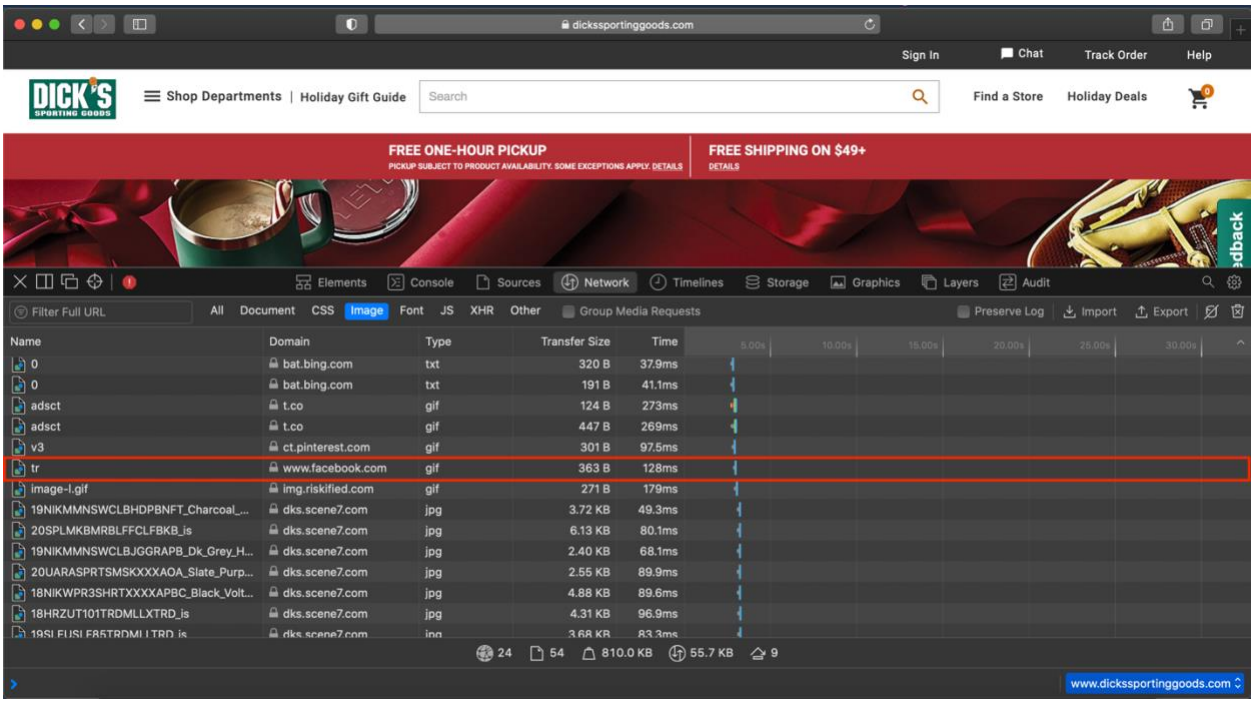
1. Open up the **Apple Safari** browser
2. Click on the **Safari** tab / **Clear History**… list item
3. Clear "all history" of related cookies and other website data

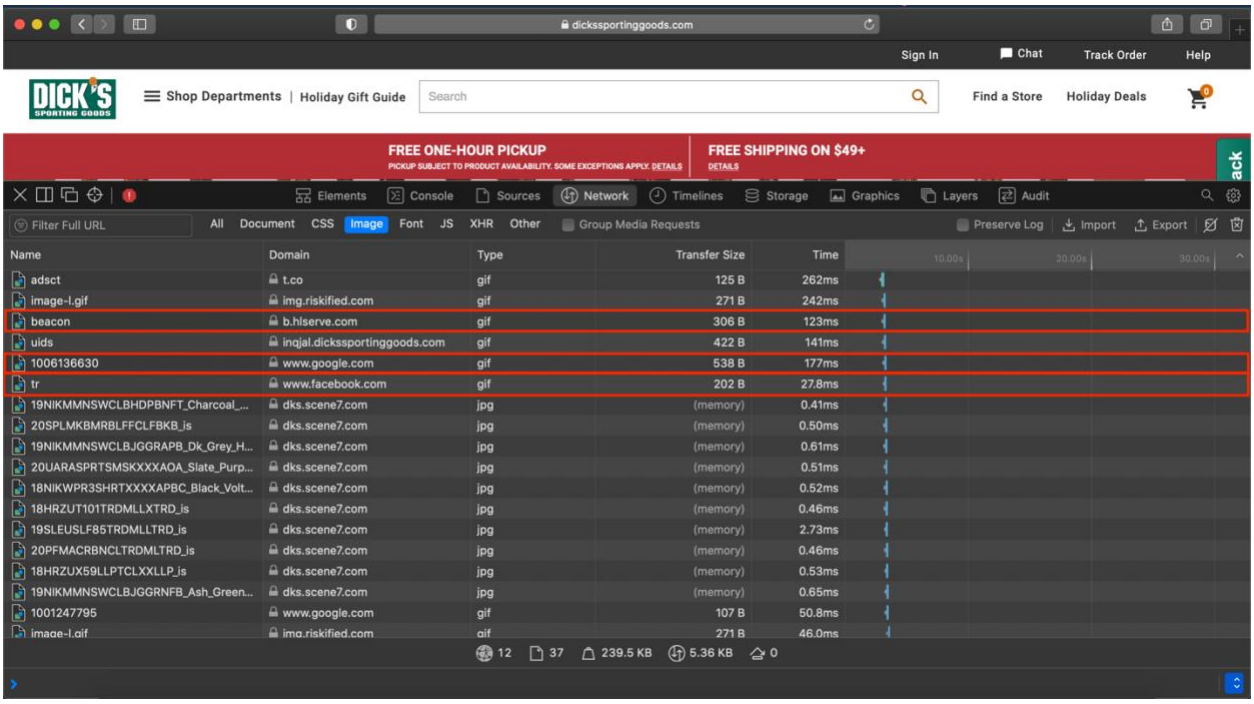


4. Go to **https://www.dicksportinggoods.com**



5. Right click on your **mouse**, select on the **Inspect Element** line item, find and select the **Network** tab, click on the **Image** filter in order to only display the image resources for **https://www.dicksportinggoods.com**
6. View result set



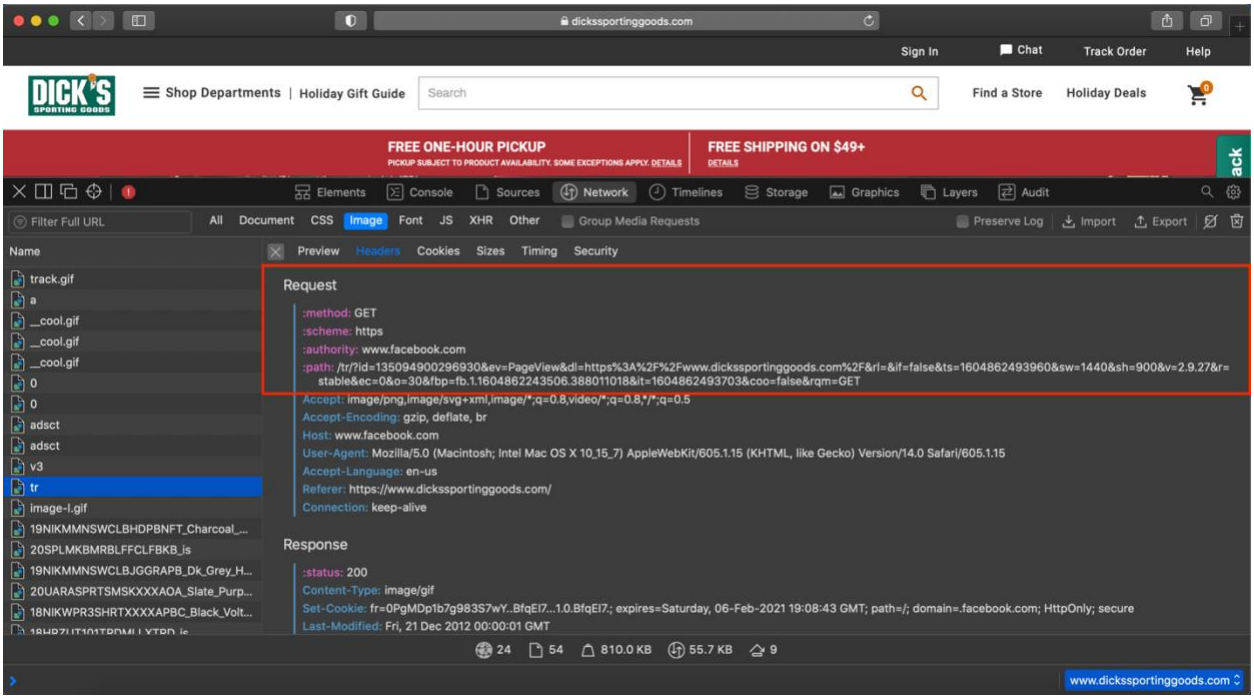


Observations:

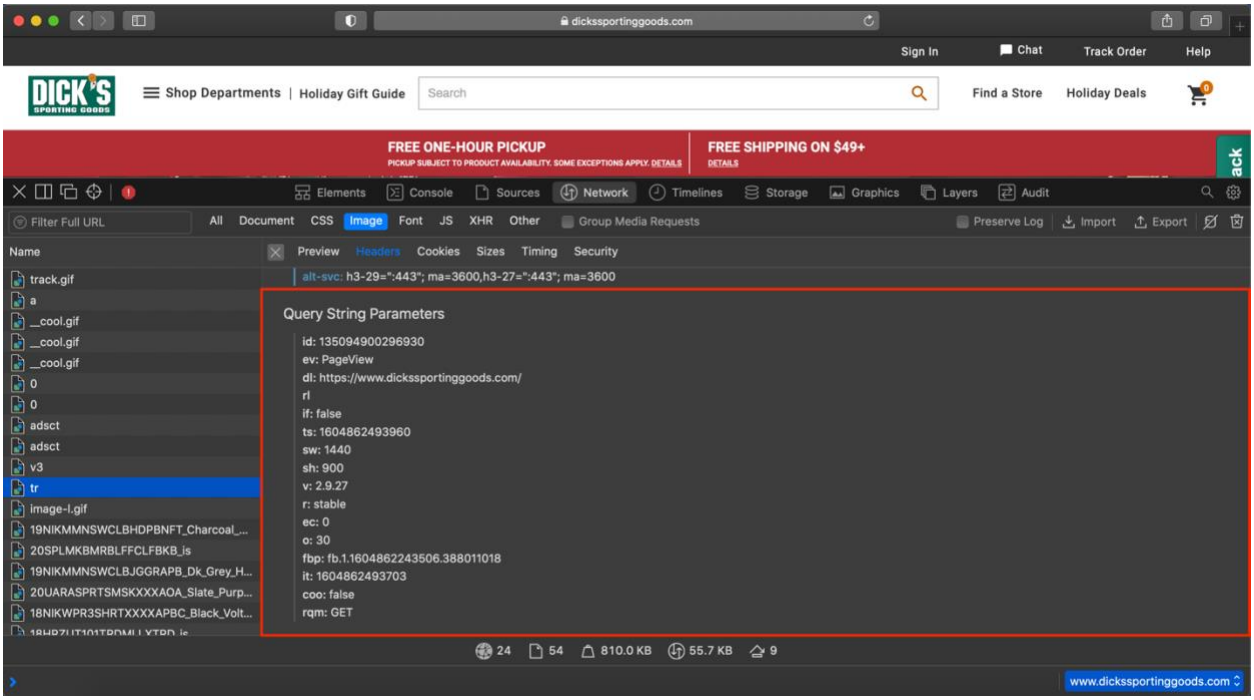
There was a total of 37 image resources that were downloaded upon initial request to https://www.dickssportinggoods.com. I suspect that at least 4 of them are web beacons. They are listed as follows:

Name:	Domain:	Type:
tr	www.facebook.com	gif
beacon	b.hlserve.com	gif
100613660	www.google.com	gif
tr	www.facebook.com	gif

If you inspect the “tr” gif request even further, you will notice that there is a bunch of data attached to the **:path** attribute of the **Request**.



What this information resolves to are called “query parameters”. They can be defined as the optional key-value pairs that appear after the question mark in a uniform resource locator (URL). In this case, query parameters provide extra data to the receiver of the **Request**.



After scanning the query parameters being sent, several of the values stand out. They are:

Key:	Abbreviation Used:	Value:
screen height	(sw)	900 (resolution)
screen width	(sw)	1440 (resolution)
domain listed?	(dl)	https://www.dickssportinggoods.com/
request made?	(rqm)	GET

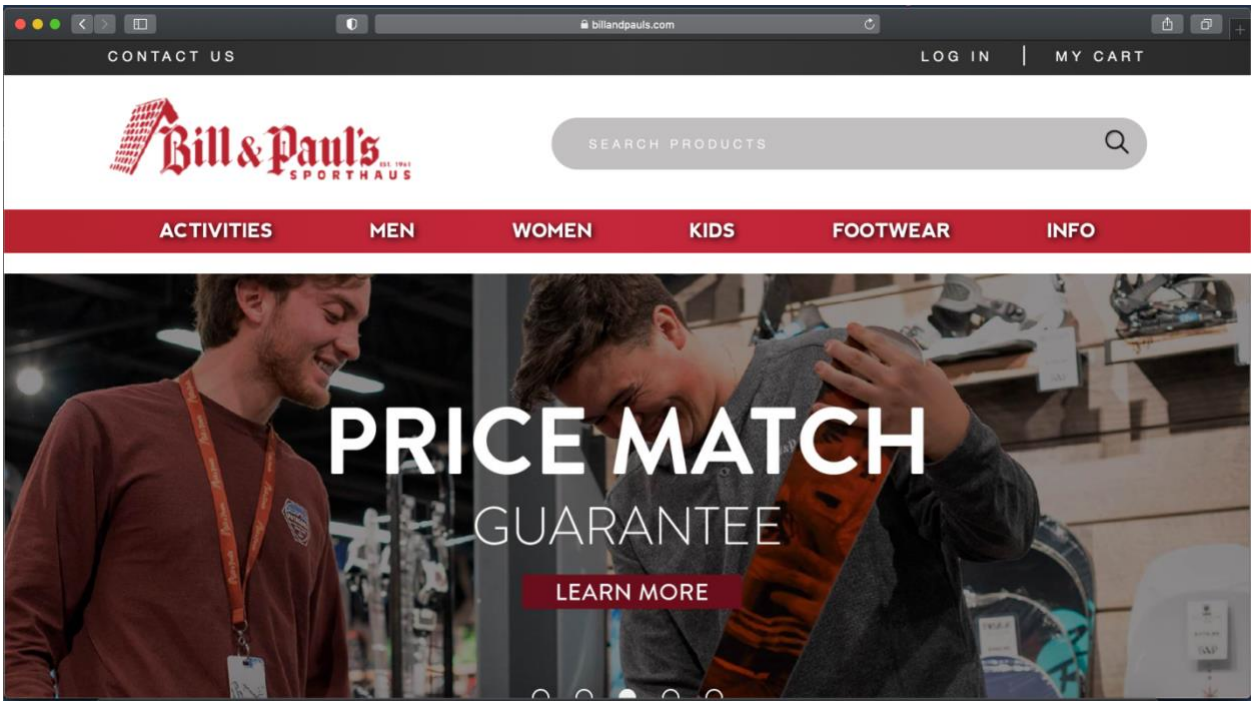
I would strongly argue that through a web beacon, Facebook is receiving the information above, amongst other data points. The company is being sent this information via query parameters. The URL used to send the data to Facebook is https://www.facebook.com/tr/?id=135094900296930&ev=PageView&dl=https%3A%2F%2Fww.dickssportinggoods.com%2F&rl=&if=false&ts=1605384905461&sw=1440&sh=900&v=2.9.28&r=stable&ec=0&o=30&fbp=fb.1.1605384064848.1961581903&it=1605384905244&coo=false&rqm=GET

Test 1.2 reproductions steps (plugin **is not** installed):

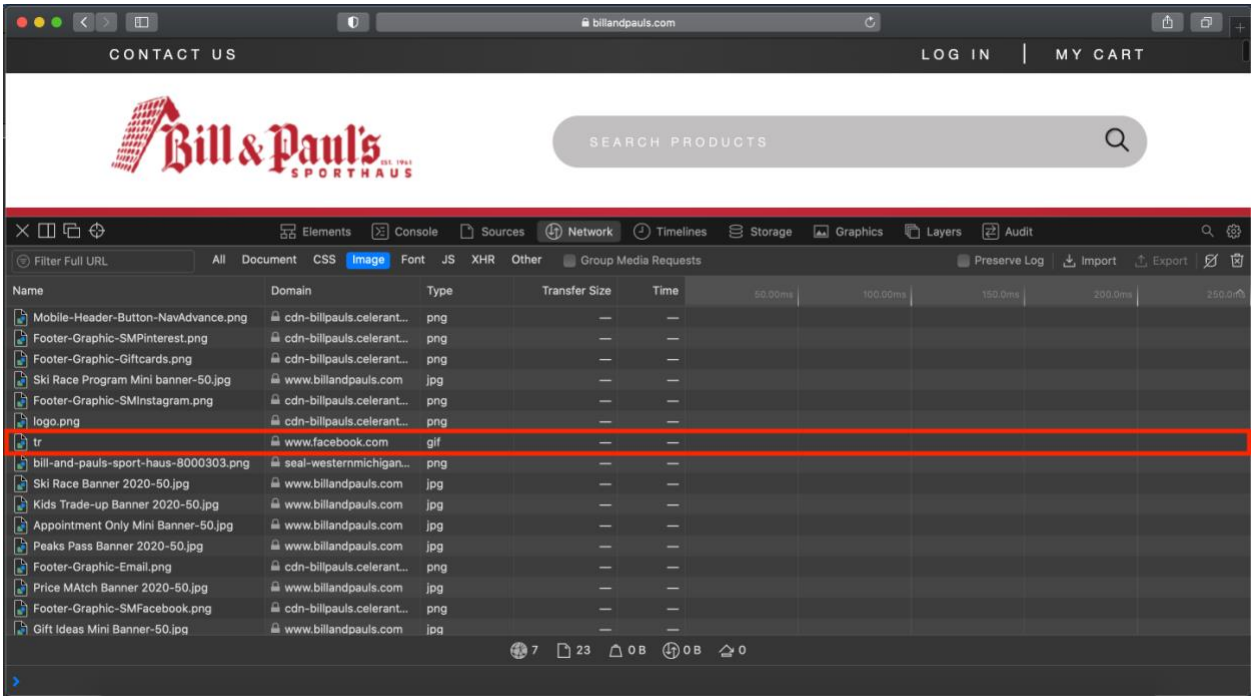
I wanted to compare **https://www.dickssportinggoods.com** with **https://www.billandpauls.com** The latter website is a local sports store located in Grand

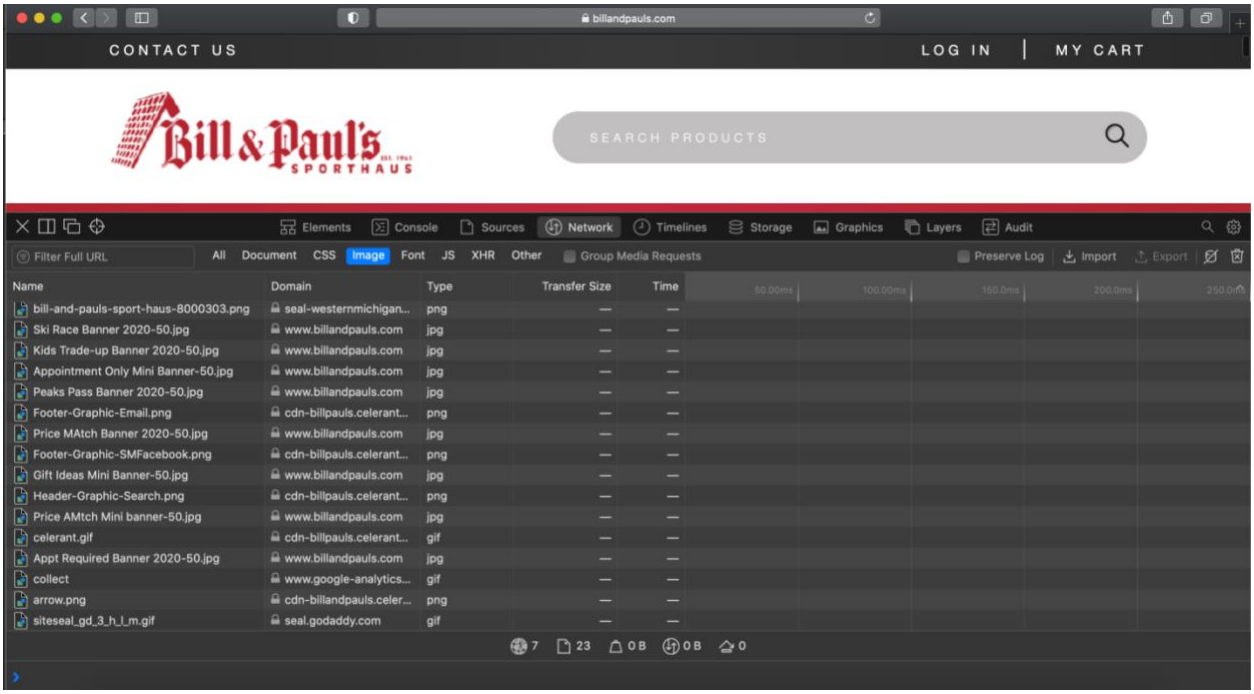
Rapids, Michigan. Their website ranks 3,121,437th out all global websites in terms of internet traffic and engagement [7].

1. Open up **Apple Safari** browser
2. Go to **https://www.billandpauls.com**



3. View image result set





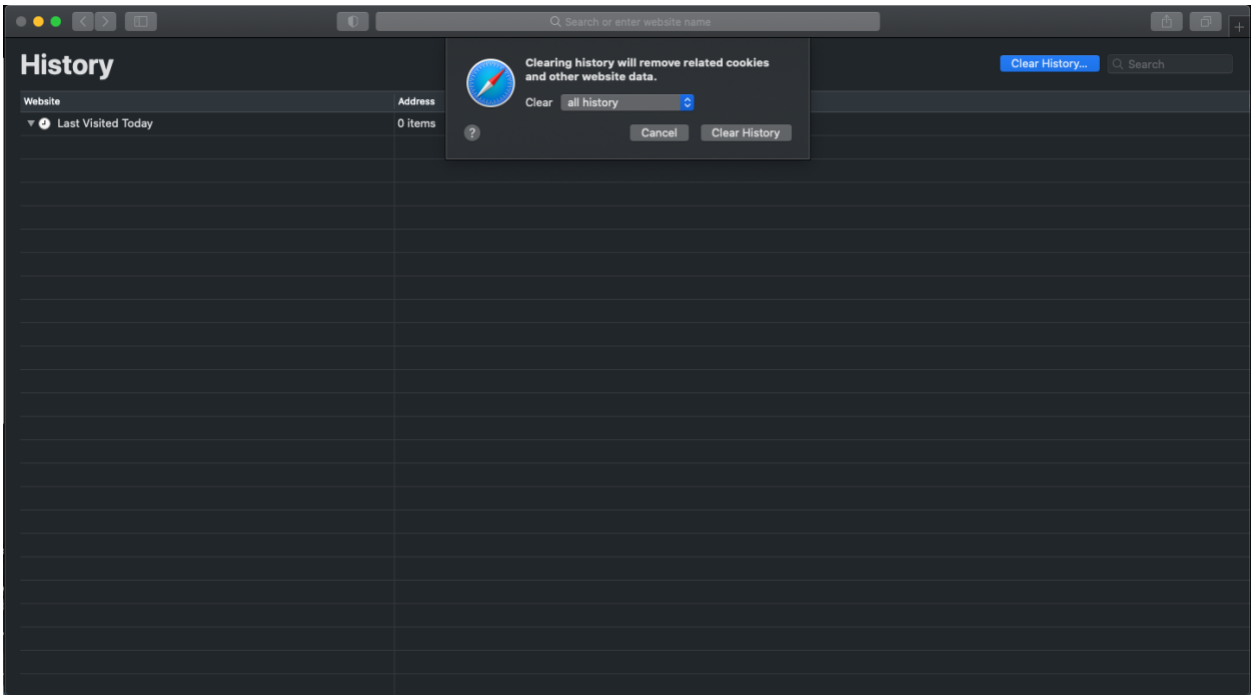
Observations:

There is 1 image that I suspect of being a web beacon.

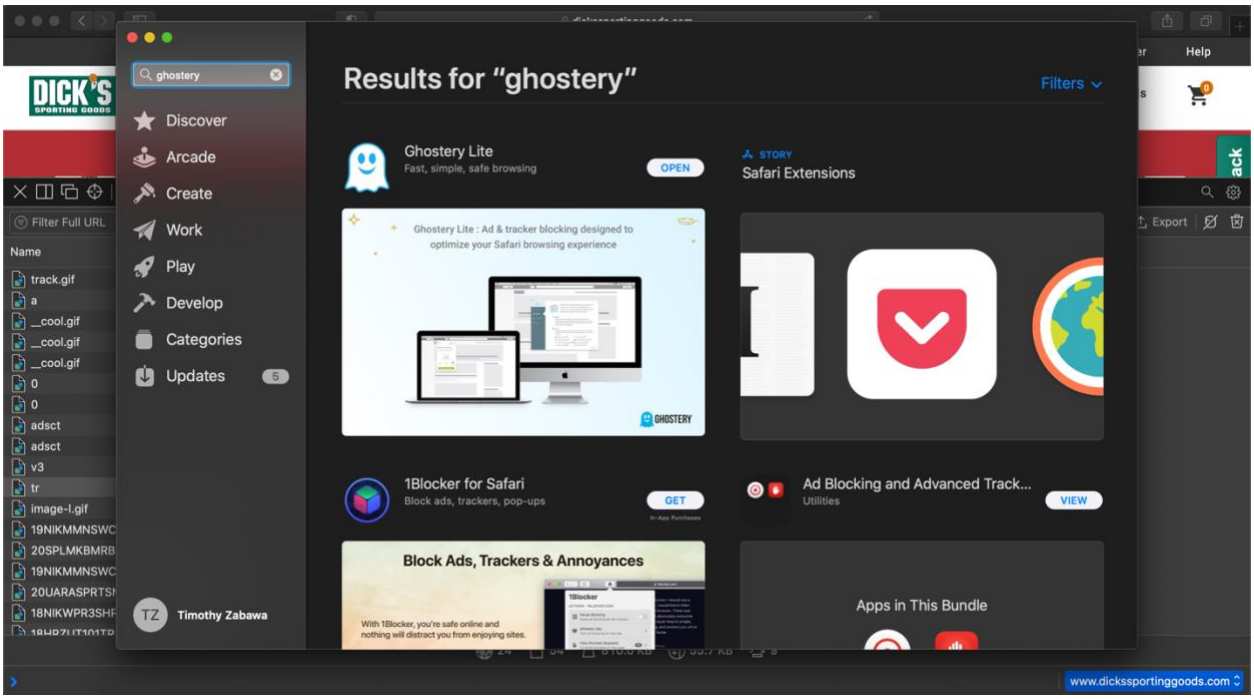
Name:	Domain:	Type:
tr	www.facebook.com	gif

Test 2 reproductions steps (plugin **is** installed):

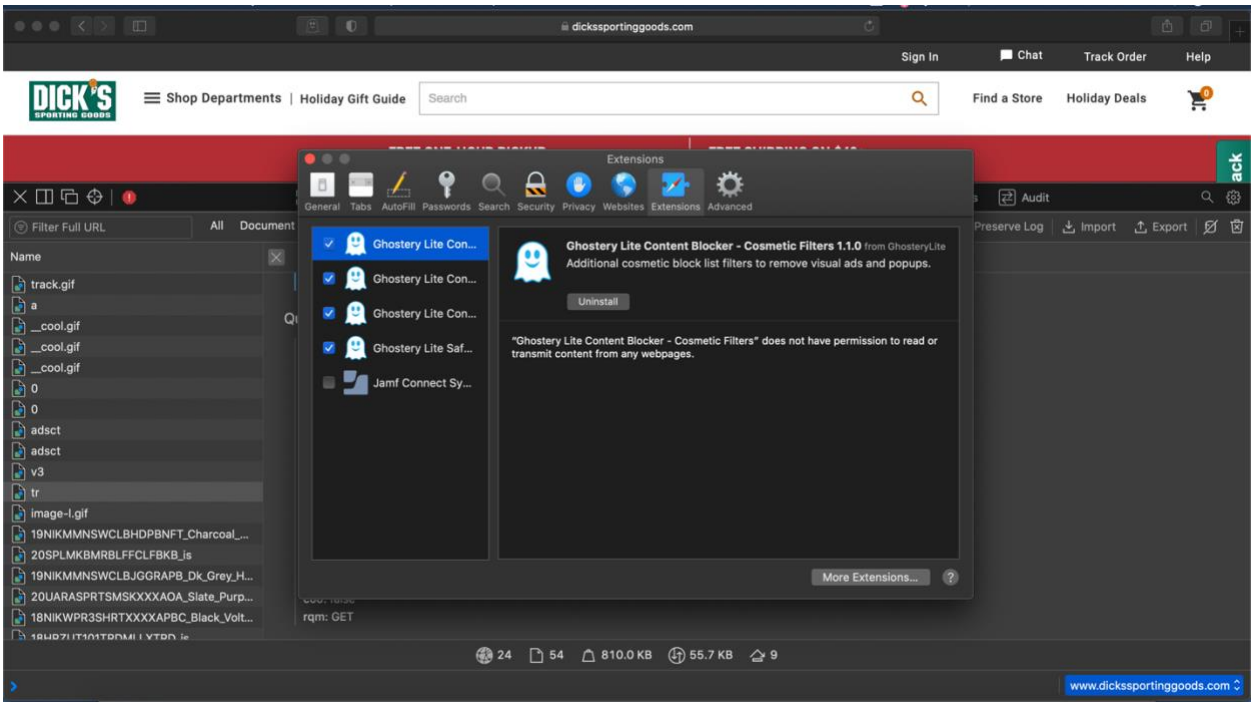
1. Open up the **Apple Safari** browser
2. Click on the **Safari** tab / **Clear History**… list item
3. Clear “all history” of related cookies and other website data



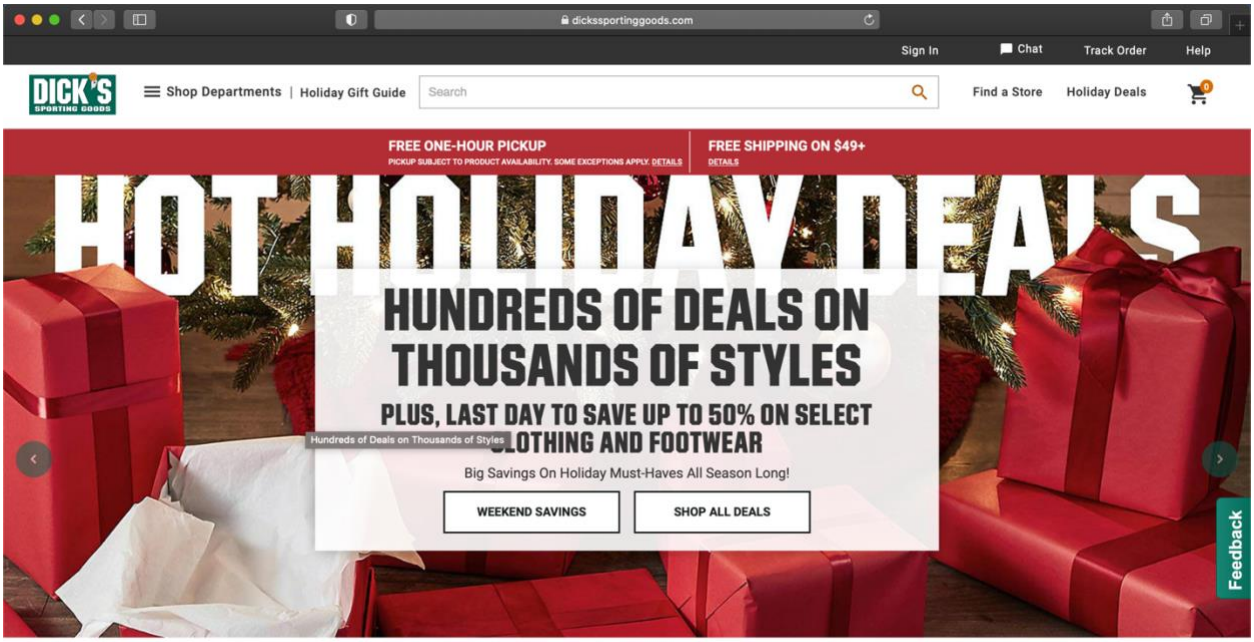
4. Click on the **Safari tab** / **Safari Extensions**… list item, search for “ghostery”, download **Ghostery Lite**



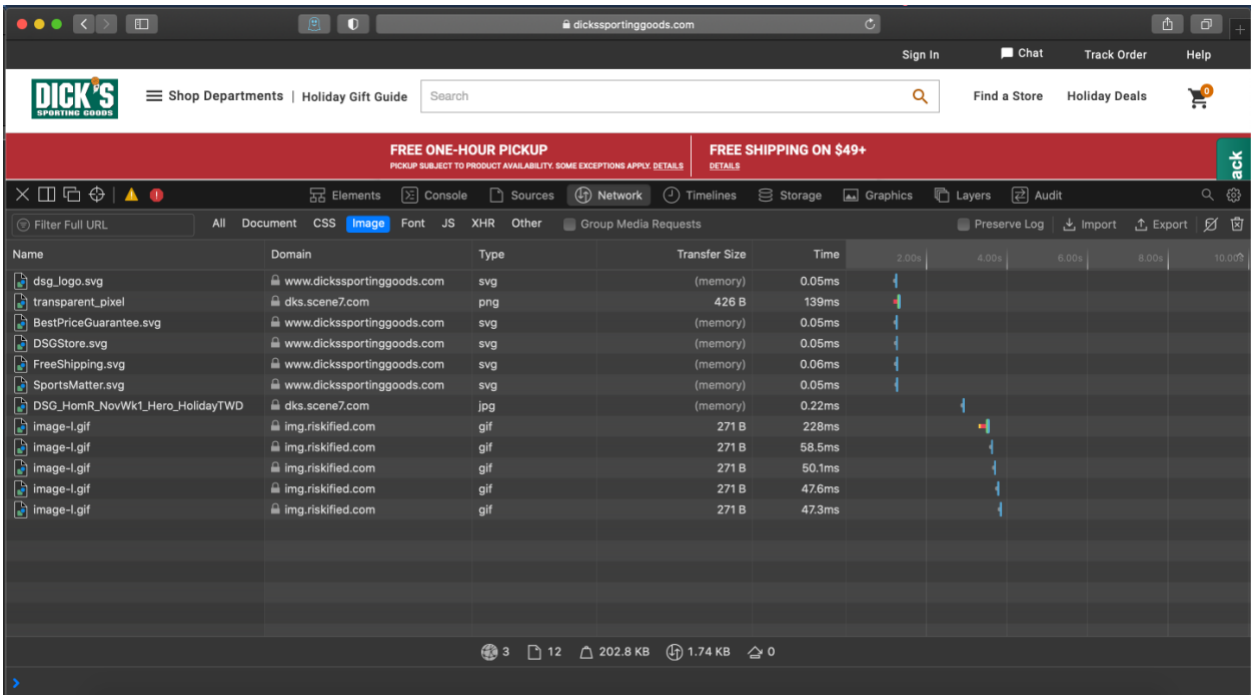
5. Click on the **Safari** tab / **Preferences**… list item, select the Extensions tab on top, activate **Ghostery Lite Content Blocker** by enabling the checkbox within each line item



6. Go to **https://www.dicksportinggoods.com**



7. Right click on your **mouse**, select on the **Inspect Element** line item, find and select the **Network** tab, click on the **Image** filter in order to only display the image resources for **https://www.dicksportinggoods.com**
8. View result set



Observations:

There was a total of 12 image resources that were downloaded upon initial request to **https://www.dickssportinggoods.com**. Ghostery seemed to have identified 25 image requests that didn't pass their "advertising network block list" filter.

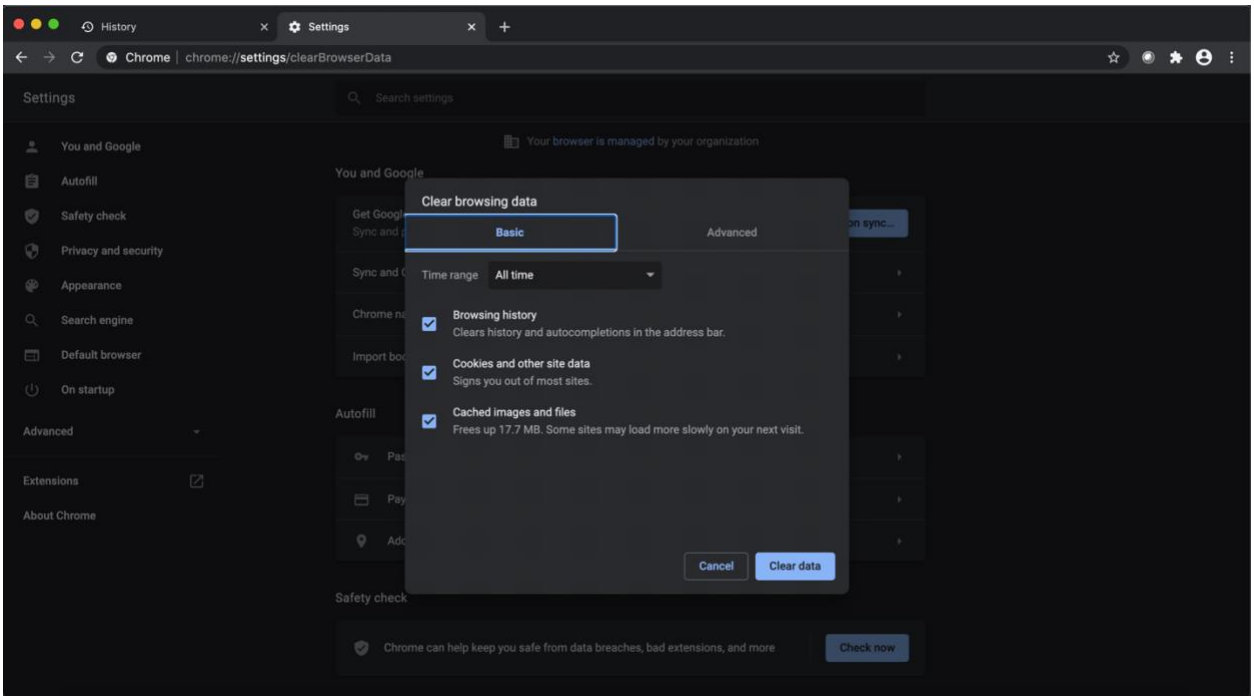
Example 2:

I carried out a test to determine how viewing a website in "incognito mode" or "private browsing" can provide protection against third-party tracking cookies.

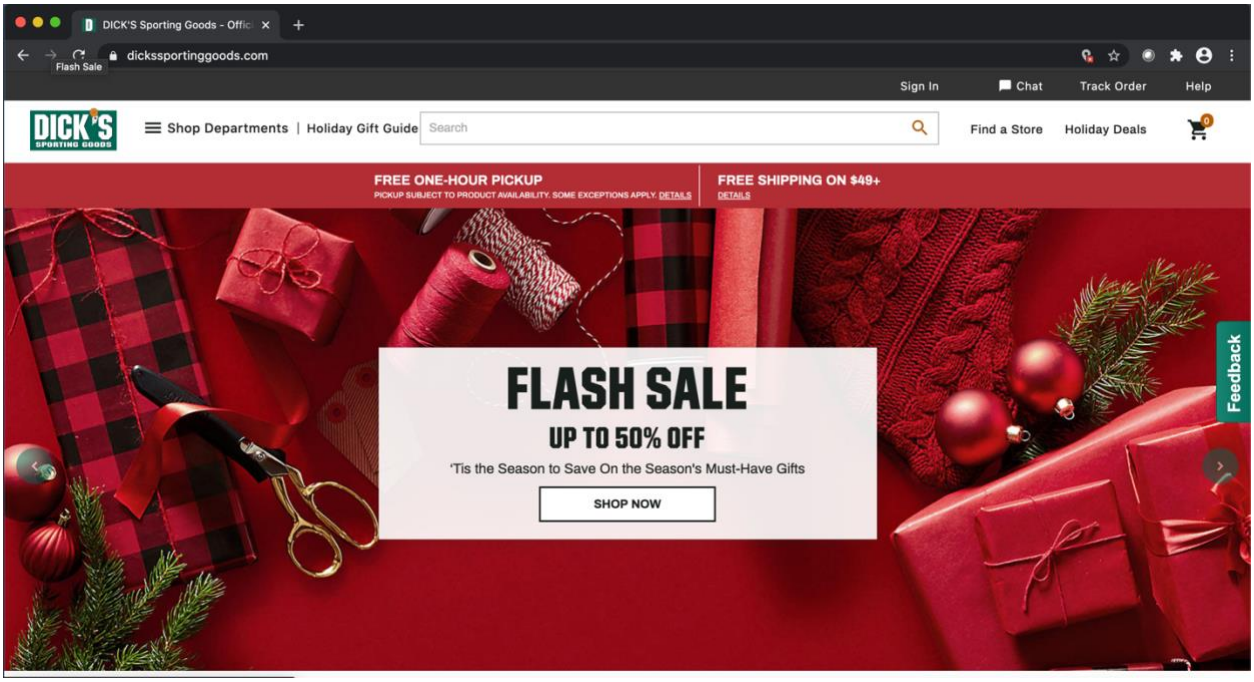
Test 1 reproduction steps (incognito mode **is not** used):

1. Open up the **Google Chrome** browser
2. Click on the **History** tab / **Show Full History**… list item
3. Click on the Clear browsing data list item on the left side of the screen

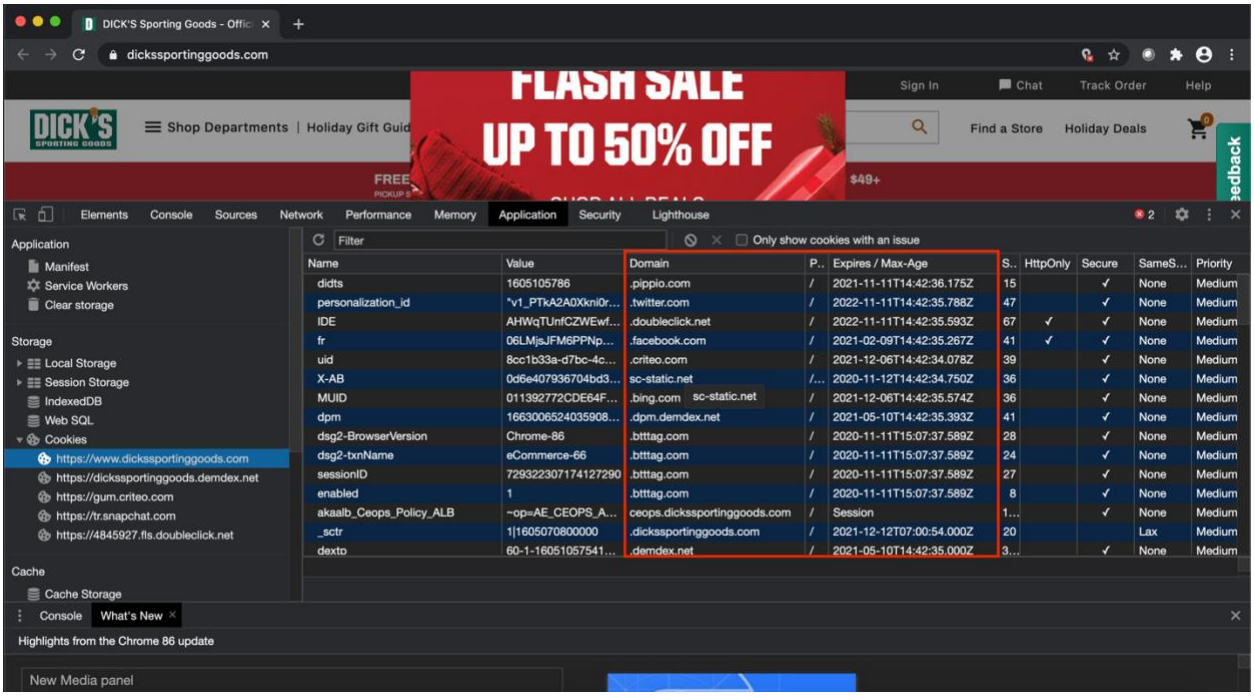
4. Select “All time” under the **Basic** tab to clear all browsing history, cookies and other site data, and cached images and files

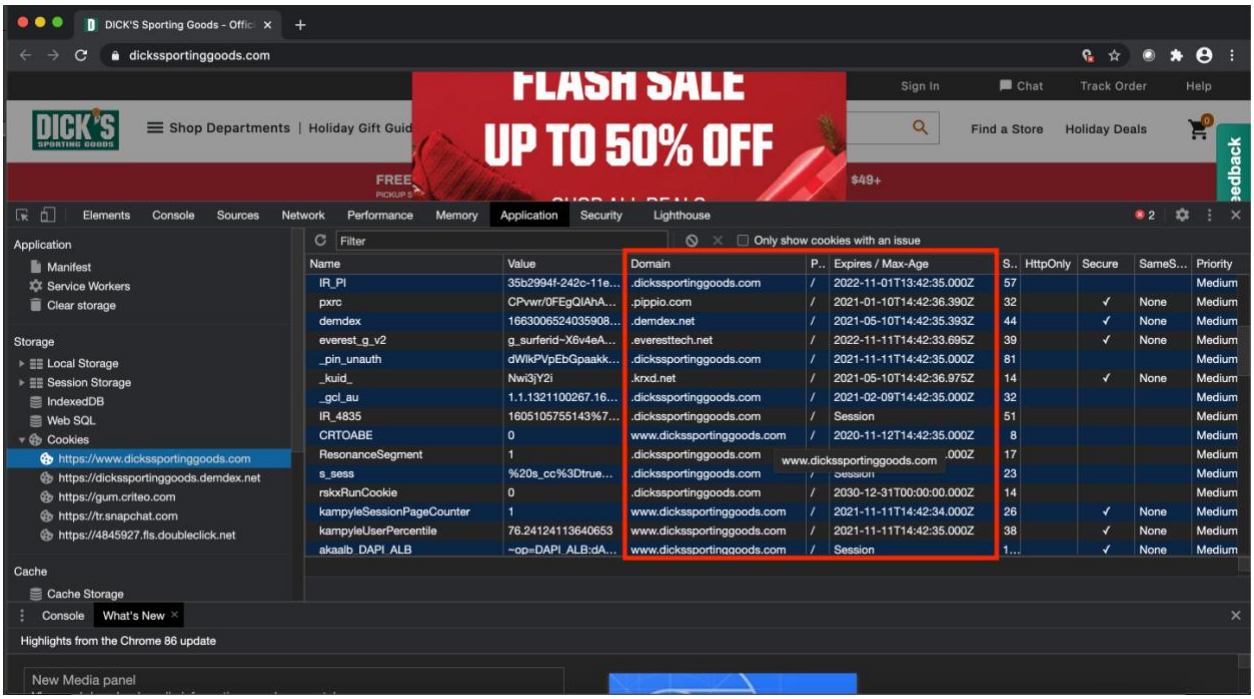


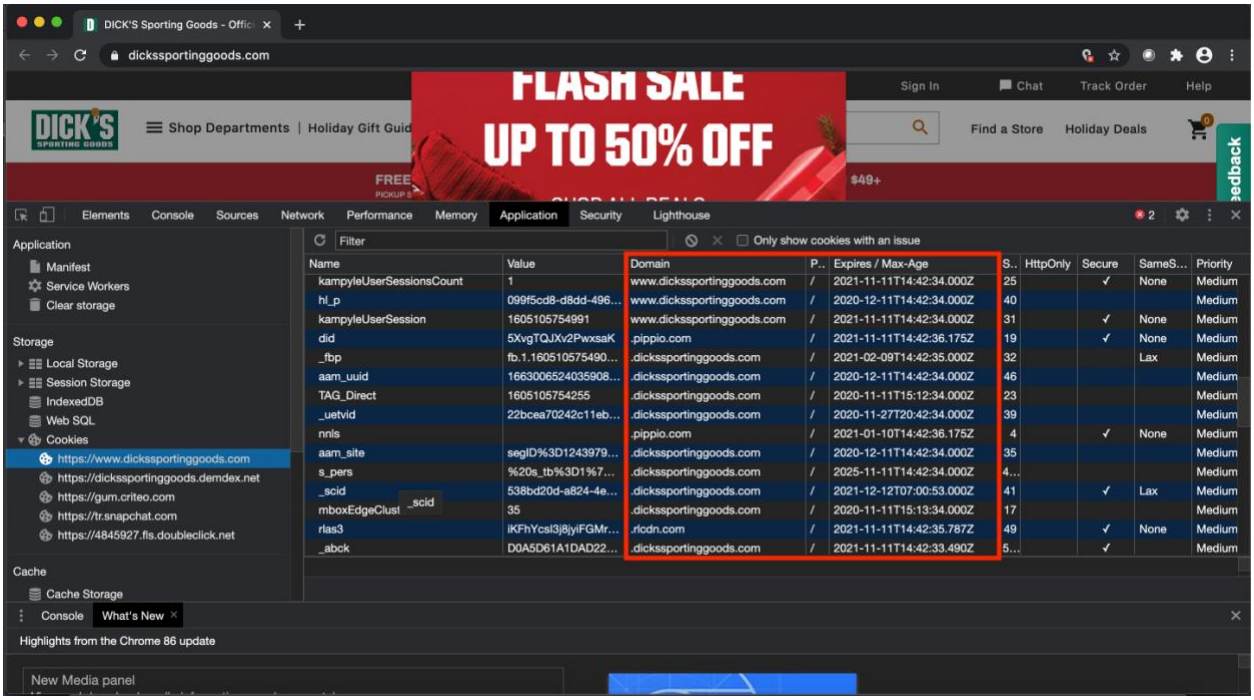
5. Go to **https://www.dicksportinggoods.com**



6. Right click on your **mouse**, select the **Inspect** line item, find and select the **Application** tab, under **Storage** / **Cookies** select **https://www.dickssportinggoods.com**
7. View result set


DICK'S Sporting Goods - Offic
dickssportinggoods.com
FLASH SALE
UP TO 50% OFF
Sign In
Chat
Track Order
Help
Shop Departments | Holiday Gift Guid
Find a Store
Holiday Deals
$49+
Elements
Console
Sources
Network
Performance
Memory
Application
Security
Lighthouse
Only show cookies with an issue
Application
Manifest
Service Workers
Clear storage
Storage
Local Storage
Session Storage
IndexedDB
Web SQL
Cookies
https://www.dickssportinggoods.com
https://dickssportinggoods.demdex.net
https://gum.criteo.com
https://tr.snapchat.com
https://4845927.fls.doubleclick.net
Cache
Cache Storage
Name
Value
Domain
Expires / Max-Age
HttpOnly
Secure
SameS...
Priority
didts
1605105786
.pippio.com
2021-11-11T14:42:36.175Z
personalization_id
"v1_PTkA2A0Xkni0r...
.twitter.com
2022-11-11T14:42:35.788Z
IDE
AHWqTUnfCZWEwf...
.doubleclick.net
2022-11-11T14:42:35.593Z
fr
06LMjsJFM6PPNp...
.facebook.com
2021-02-09T14:42:35.267Z
uid
8cc1b33a-d7bc-4c...
.criteo.com
2021-12-06T14:42:34.078Z
X-AB
0d6e407936704bd3...
sc-static.net
2020-11-12T14:42:34.750Z
MUID
011392772CDE64F...
.bing.com
sc-static.net
2021-12-06T14:42:35.574Z
dpm
1663006524035908...
.dpm.demdex.net
2021-05-10T14:42:35.393Z
dsg2-BrowserVersion
Chrome-86
.btttag.com
2020-11-11T15:07:37.589Z
dsg2-txnName
eCommerce-66
.btttag.com
2020-11-11T15:07:37.589Z
sessionID
729322307174127290
.btttag.com
2020-11-11T15:07:37.589Z
enabled
1
.btttag.com
2020-11-11T15:07:37.589Z
akaalb_Ceops_Policy_ALB
~op=AE_CEOPS_A...
ceops.dickssportinggoods.com
Session
_sctr
1|1605070800000
.dickssportinggoods.com
2021-12-12T07:00:54.000Z
Lax
dextp
60-1-16051057541...
.demdex.net
2021-05-10T14:42:35.000Z
Console
What's New
Highlights from the Chrome 86 update
New Media panel


DICK'S Sporting Goods - Offic
dickssportinggoods.com
FLASH SALE
UP TO 50% OFF
Sign In
Chat
Track Order
Help
Shop Departments | Holiday Gift Guid
Find a Store
Holiday Deals
$49+
Elements
Console
Sources
Network
Performance
Memory
Application
Security
Lighthouse
Only show cookies with an issue
Application
Manifest
Service Workers
Clear storage
Storage
Local Storage
Session Storage
IndexedDB
Web SQL
Cookies
https://www.dickssportinggoods.com
https://dickssportinggoods.demdex.net
https://gum.criteo.com
https://tr.snapchat.com
https://4845927.fls.doubleclick.net
Cache
Cache Storage
Name
Value
Domain
Expires / Max-Age
HttpOnly
Secure
SameS...
Priority
IR_PI
35b2994f-242c-11e...
.dickssportinggoods.com
2022-11-01T13:42:35.000Z
pxrc
CPvwr/0FEgQIAhA...
.pippio.com
2021-01-10T14:42:36.390Z
demdex
1663006524035908...
.demdex.net
2021-05-10T14:42:35.393Z
everest_g_v2
g_surferid~X6v4eA...
.everesttech.net
2022-11-11T14:42:33.695Z
_pin_unauth
dWlkPVpEbGpaakk...
.dickssportinggoods.com
2021-11-11T14:42:35.000Z
kuid
Nwi3jY2i
.krxd.net
2021-05-10T14:42:36.975Z
_gcl_au
1.1.1321100267.16...
.dickssportinggoods.com
2021-02-09T14:42:35.000Z
IR_4835
1605105755143%7...
.dickssportinggoods.com
Session
CRTOABE
0
www.dickssportinggoods.com
2020-11-12T14:42:35.000Z
ResonanceSegment
1
.dickssportinggoods.com
www.dickssportinggoods.com
s_sess
%20s_cc%3Dtrue...
.dickssportinggoods.com
rskxRunCookie
0
.dickssportinggoods.com
2030-12-31T00:00:00.000Z
kampyleSessionPageCounter
1
www.dickssportinggoods.com
2021-11-11T14:42:34.000Z
kampyleUserPercentile
76.24124113640653
www.dickssportinggoods.com
2021-11-11T14:42:35.000Z
akaalb_DAPI_ALB
~op=DAPI_ALB:dA...
www.dickssportinggoods.com
Session
Console
What's New
Highlights from the Chrome 86 update
New Media panel



Observations:

There was a total of 71 cookies attached to the browser after **https://www.dickssportinggoods.com** was loaded. Of that number, 56 were requested from an external source that shared the same domain (dickssportinggoods.com). That leaves 15 cookies that came from third-party origins. I suspect that at least 5 of them are tracking cookies. They are listed as follows:

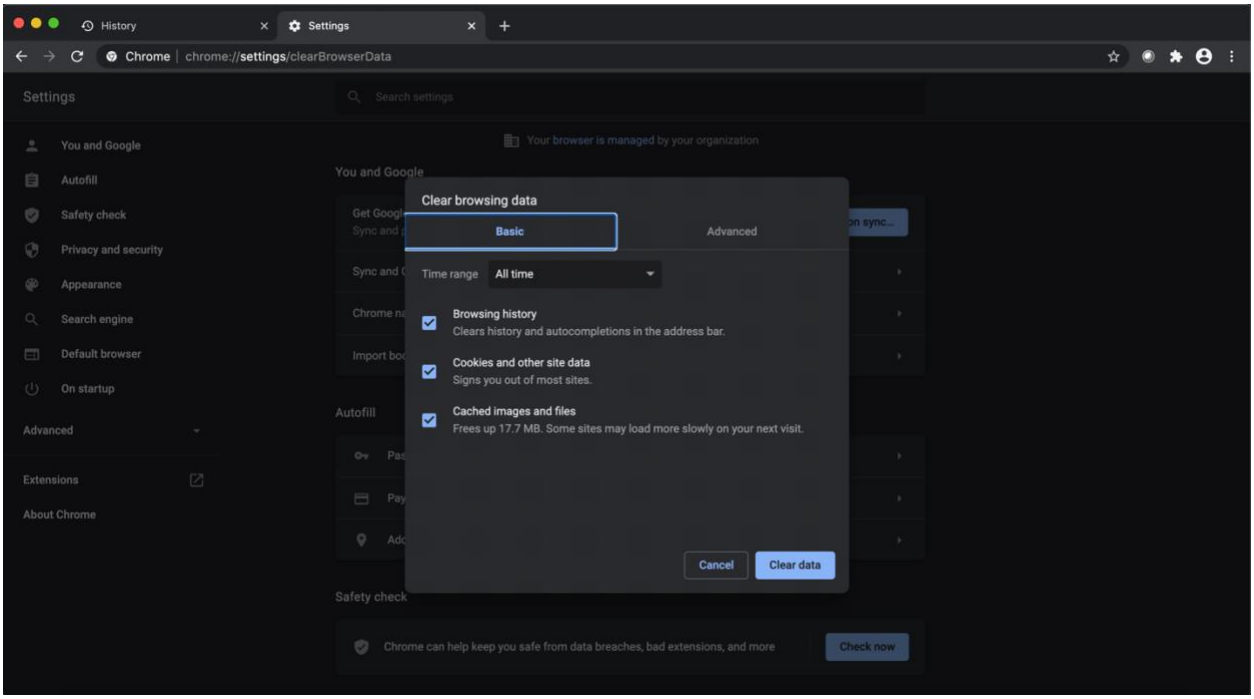
Name:	Domain:	Expires/Max-Age:
personalization_id	.twitter.com	2022-11-11T…
1P_JAR	.google.com	2020-12-11T…
NID	.google.com	2021-05-13T…
fr	.facebook.com	2021-02-09T…
IDE	.doubleclick.net	2021-11-11T…

The cookies above are coming from companies such as Twitter, Google, and Facebook. Their expiration date ranges from the beginning of year 2021 through the 2022. While they are on the browser, information about the user can be added to the cookie and eventually shared with the

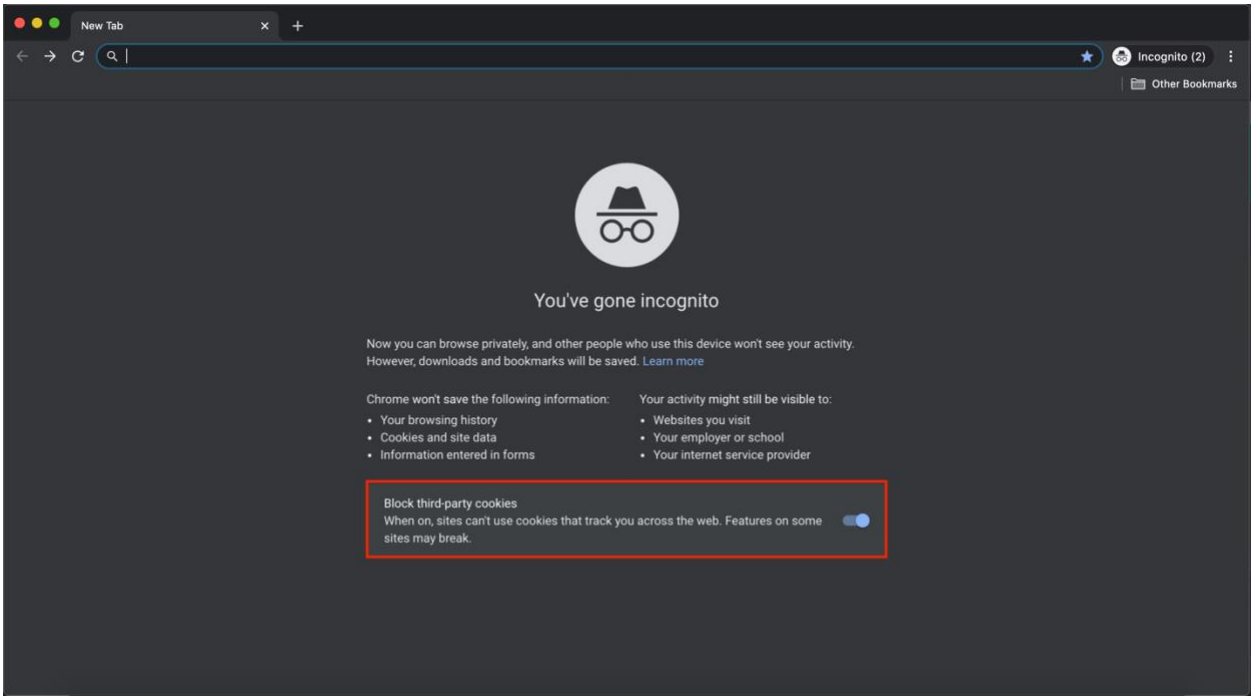
respectful authors. Because of the domain and expiration date of these cookies, I argue that they are third-party tracking cookies.

Test 2 reproduction steps (incognito mode **is** used):

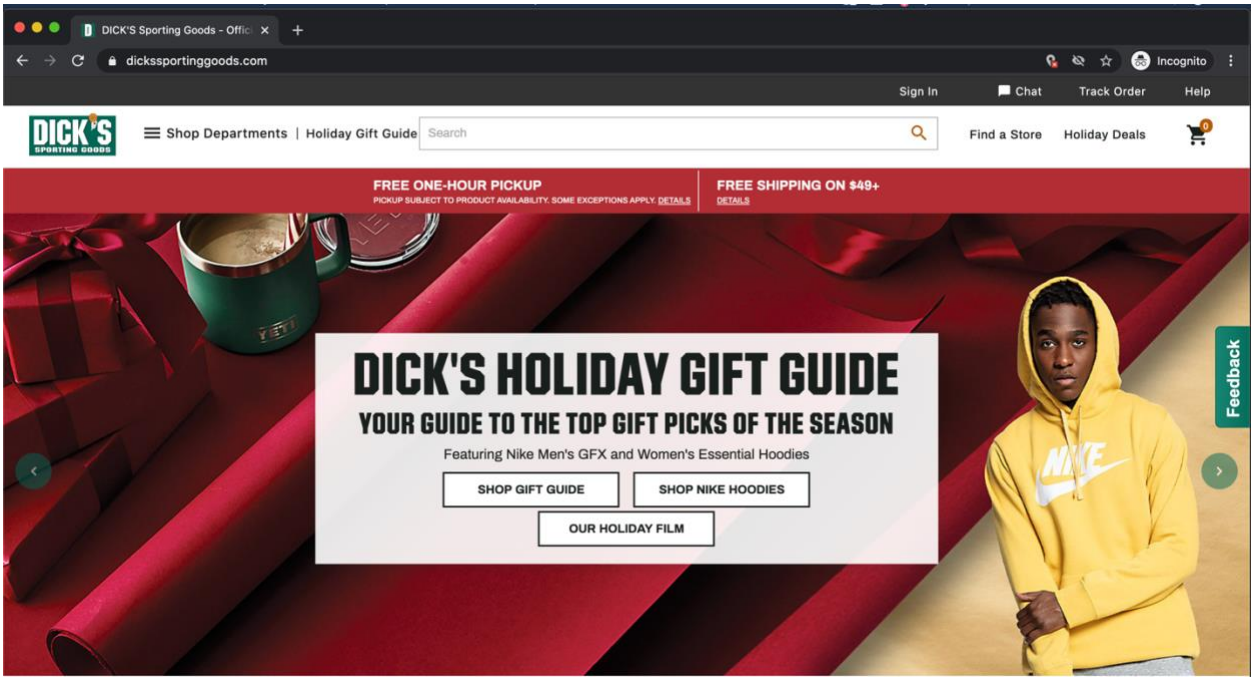
1. Open up the **Google Chrome** browser
2. Click on the **History** tab / **Show Full History**… list item
3. Click on the Clear browsing data list item on the left side of the screen
4. Select “All time” under the Basic tab to clear all browsing history, cookies and other site data, and cached images and files



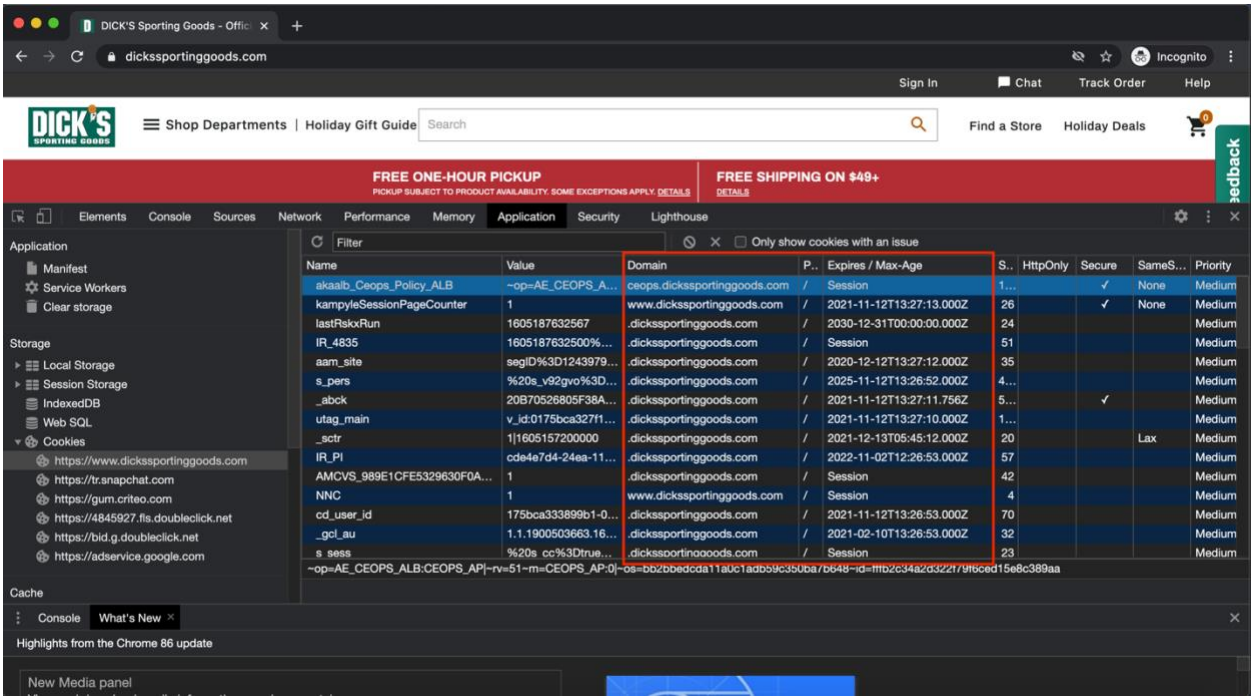
5. From the main browser, click on the **File** tab / **New Incognito Window** list item

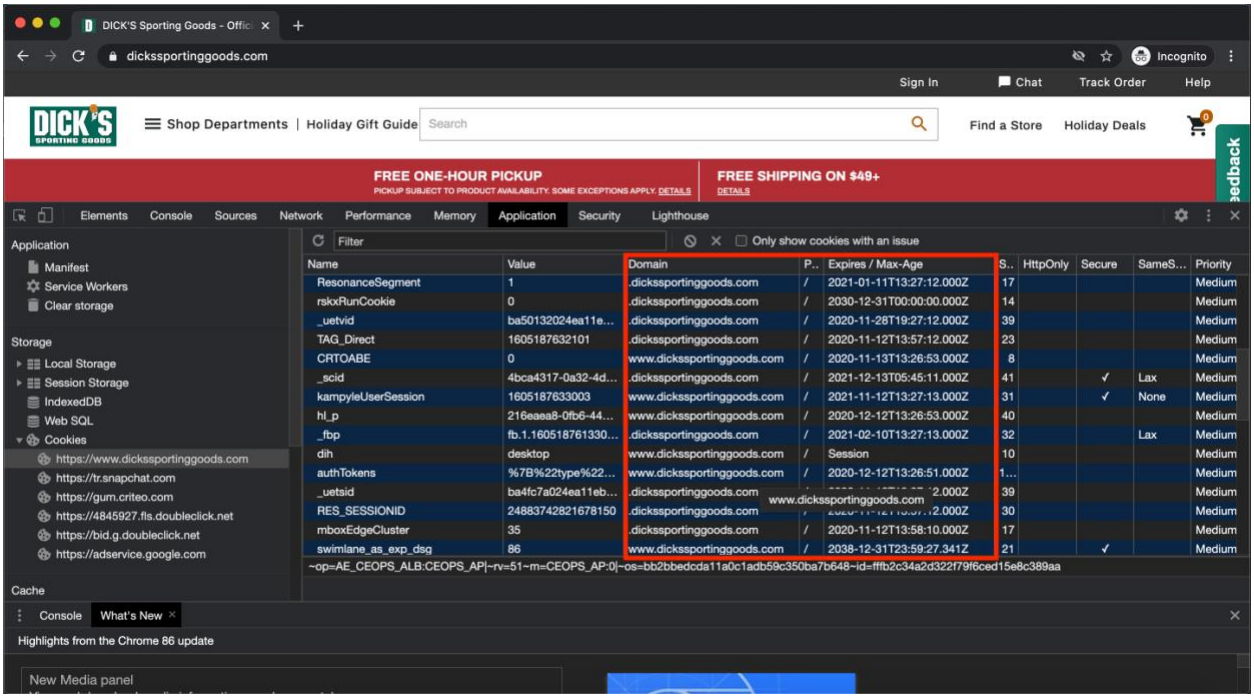


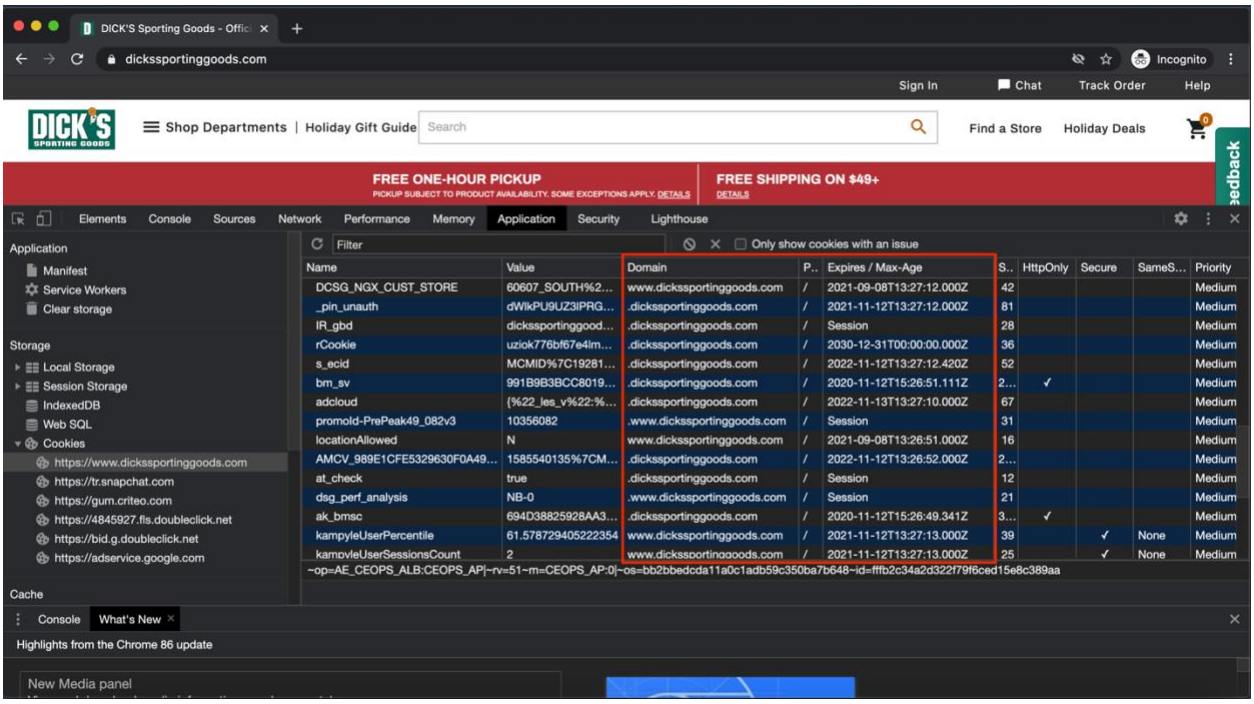
6. Go to **https://www.dicksportinggoods.com**

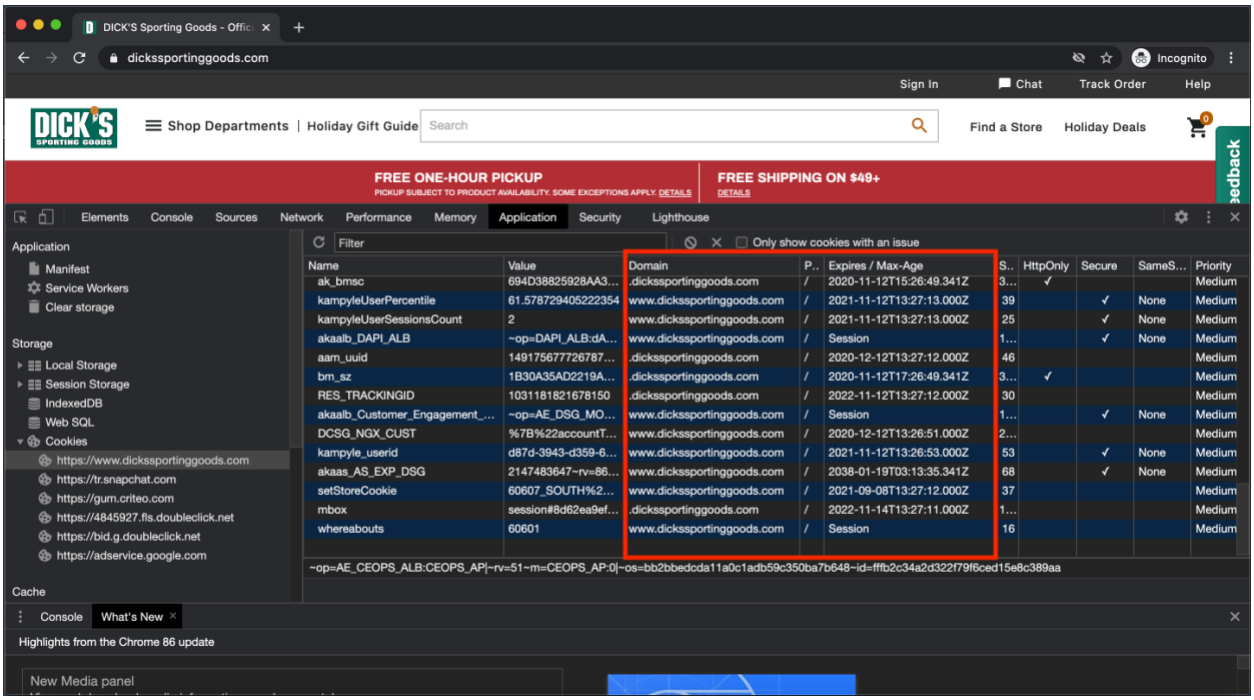


7. Right click on your **mouse**, select the **Inspect** line item, find and select the **Application** tab, under **Storage** / **Cookies** select **https://www.dickssportinggoods.com**
8. View result set


DICK'S Sporting Goods - Offic
dickssportinggoods.com
Incognito
Sign In
Chat
Track Order
Help
Shop Departments | Holiday Gift Guide
Search
Find a Store
Holiday Deals
FREE ONE-HOUR PICKUP
PICKUP SUBJECT TO PRODUCT AVAILABILITY. SOME EXCEPTIONS APPLY. DETAILS
FREE SHIPPING ON $49+
DETAILS
Elements
Console
Sources
Network
Performance
Memory
Application
Security
Lighthouse
Filter
Only show cookies with an issue
Application
Manifest
Service Workers
Clear storage
Storage
Local Storage
Session Storage
IndexedDB
Web SQL
Cookies
https://www.dickssportinggoods.com
https://tr.snapchat.com
https://gum.criteo.com
https://4845927.fls.doubleclick.net
https://bid.g.doubleclick.net
https://adservice.google.com
Cache
Name
Value
Domain
P..
Expires / Max-Age
S..
HttpOnly
Secure
SameS...
Priority
akaalb_Ceops_Policy_ALB
~op=AE_CEOPS_A...
ceops.dickssportinggoods.com
/
Session
1...
✓
None
Medium
kampyleSessionPageCounter
1
www.dickssportinggoods.com
/
2021-11-12T13:27:13.000Z
26
✓
None
Medium
lastRskxRun
1605187632567
.dickssportinggoods.com
/
2030-12-31T00:00:00.000Z
24
Medium
IR_4835
16051876325000...
.dickssportinggoods.com
/
Session
51
Medium
aam_site
segID%3D1243979...
.dickssportinggoods.com
/
2020-12-12T13:27:12.000Z
35
Medium
s_pers
%20s_v92gvo%3D...
.dickssportinggoods.com
/
2025-11-12T13:26:52.000Z
4...
Medium
_abck
20B70526805F38A...
.dickssportinggoods.com
/
2021-11-12T13:27:11.756Z
5...
✓
Medium
utag_main
v_id:0175bca327f1...
.dickssportinggoods.com
/
2021-11-12T13:27:10.000Z
1...
Medium
_sctr
1|1605157200000
.dickssportinggoods.com
/
2021-12-13T05:45:12.000Z
20
Lax
Medium
IR_PI
cde4e7d4-24ea-11...
.dickssportinggoods.com
/
2022-11-02T12:26:53.000Z
57
Medium
AMCVS_989E1CFE5329630F0A...
1
.dickssportinggoods.com
/
Session
42
Medium
NNC
1
www.dickssportinggoods.com
/
Session
4
Medium
cd_user_id
175bca333899b1-0...
.dickssportinggoods.com
/
2021-11-12T13:26:53.000Z
70
Medium
_gcl_au
1.1.1900503663.16...
.dickssportinggoods.com
/
2021-02-10T13:26:53.000Z
32
Medium
s_sess
%20s_cc%3Dtrue...
.dickssportinggoods.com
/
Session
23
Medium
~op=AE_CEOPS_ALB:CEOPS_AP|~rv=51~m=CEOPS_AP:0|~os=bb2bbedcda11a0c1adb59c350ba7b648~id=fffb2c34a2d322f79f6ced15e8c389aa
Console
What's New
Highlights from the Chrome 86 update
New Media panel


DICK'S Sporting Goods - Offic
dickssportinggoods.com
Incognito
Sign In
Chat
Track Order
Help
Shop Departments | Holiday Gift Guide
Search
Find a Store
Holiday Deals
FREE ONE-HOUR PICKUP
PICKUP SUBJECT TO PRODUCT AVAILABILITY. SOME EXCEPTIONS APPLY. DETAILS
FREE SHIPPING ON $49+
DETAILS
Elements
Console
Sources
Network
Performance
Memory
Application
Security
Lighthouse
Filter
Only show cookies with an issue
Application
Manifest
Service Workers
Clear storage
Storage
Local Storage
Session Storage
IndexedDB
Web SQL
Cookies
https://www.dickssportinggoods.com
https://tr.snapchat.com
https://gum.criteo.com
https://4845927.fls.doubleclick.net
https://bid.g.doubleclick.net
https://adservice.google.com
Cache
Name
Value
Domain
P..
Expires / Max-Age
S..
HttpOnly
Secure
SameS...
Priority
ResonanceSegment
1
.dickssportinggoods.com
/
2021-01-11T13:27:12.000Z
17
Medium
rskxRunCookie
0
.dickssportinggoods.com
/
2030-12-31T00:00:00.000Z
14
Medium
_uetvid
ba50132024ea11e...
.dickssportinggoods.com
/
2020-11-28T19:27:12.000Z
39
Medium
TAG_Direct
1605187632101
.dickssportinggoods.com
/
2020-11-12T13:57:12.000Z
23
Medium
CRTOABE
0
www.dickssportinggoods.com
/
2020-11-13T13:26:53.000Z
8
Medium
_scid
4bca4317-0a32-4d...
.dickssportinggoods.com
/
2021-12-13T05:45:11.000Z
41
✓
Lax
Medium
kampyleUserSession
1605187633003
www.dickssportinggoods.com
/
2021-11-12T13:27:13.000Z
31
✓
None
Medium
hl_p
216eaea8-0fb6-44...
www.dickssportinggoods.com
/
2020-12-12T13:26:53.000Z
40
Medium
_fbp
fb.1.160518761330...
.dickssportinggoods.com
/
2021-02-10T13:27:13.000Z
32
Lax
Medium
dih
desktop
www.dickssportinggoods.com
/
Session
10
Medium
authTokens
%7B%22type%22...
www.dickssportinggoods.com
/
2020-12-12T13:26:51.000Z
1...
Medium
_uetsid
ba4fc7a024ea11eb...
.dickssportinggoods.com
www.dickssportinggoods.com
2.000Z
39
Medium
RES_SESSIONID
2488374282167815O
.dickssportinggoods.com
/
2.000Z
30
Medium
mboxEdgeCluster
35
.dickssportinggoods.com
/
2020-11-12T13:58:10.000Z
17
Medium
swimlane_as_exp_dsg
86
www.dickssportinggoods.com
/
2038-12-31T23:59:27.341Z
21
✓
Medium
~op=AE_CEOPS_ALB:CEOPS_AP|~rv=51~m=CEOPS_AP:0|~os=bb2bbedcda11a0c1adb59c350ba7b648~id=fffb2c34a2d322f79f6ced15e8c389aa
Console
What's New
Highlights from the Chrome 86 update
New Media panel





Observations:

After viewing the same website in incognito mode, the only cookies that were attached to the browser where the 56 that shared the same domain (dickssportinggoods.com). It should be highlighted that web tracking may still be taking place. In this case, Dicks Sporting Goods may be trying to track the fact that you visited the website. They also may be monitoring your actions as

you shop. However, by viewing the web page in incognito mode, companies such as Twitter, Google, and Facebook are not able to place third-party tracking cookies on your browser.

Overall, these tests show that defensive measures, such as browser plugins and surfing the web in incognito mode, can have positive effects on one's online privacy. Ghostery successfully targeted and prevented certain web beacons that didn't pass their "advertising network block list" filter. Google's incognito mode disallowed third-party tracking cookies from attaching themselves to a browser. Comparing **https://www.dickssportinggoods.com** with **https://www.billandpauls.com** highlighted the fact that companies like Facebook are very effective at their tracking capabilities. It is important to know that some of these defenses can contain side effects that will negatively affect a user's internet experience. Tradeoffs will have to be accepted as one tries to utilize the web more privately.

5. Why consumder data is sought after:

As the 21st century goes on, business will continue to evolve. It is safe to say that at least part of this evolution will involve the Internet. The application continues to offer an increasing number of products and services. As users spend more time on the web, businesses want to know how it is being utilized. They wish to use this data for a wide range of reasons, including audience targeting, selling user information as a product, and malicious intent.

It is widely known that advertisers use online data to try and make better decisions. A lot of money is wasted on resources put into products that customers don't actually want. Businesses need information on what is happening internally and externally within their organization [8]. Consumer data can be used to create effective company models that have a thorough understanding of what products attract the most clientele. Audience targeting has become more popular as online information has become cheaper. Big data is no longer siloed within huge corporations with large financial resources. Small businesses now have a substantial amount of exposure to big data products [9].

Prospective customers are not the only end goal for raw user information. In fact, some companies actually re-sell consumed data as a product in itself. These businesses are known as data brokers [10]. Services include aggregated collections of bare web data.

Although consumers might be bothered to know that their information is being used in these ways, it is legal. Unfortunately, there also exists malicious ways data can be employed. Thieves can use information shared online to steal identities or personal possessions. Users should be mindful of the risks they put themselves in depending on what data is exposed. All in all, the ways consumer information can be taken advantage of is growing by the day. An important point to remember is that unlike times before the Internet, there is no expiration date on appropriately stored online data. Once gathered, companies have the ability to exploit the information indefinitely.

6. Conclusion:

It is important for users to understand that surfing the web is not truly "free". What is given in exchange for accessing the Internet is information about oneself. Businesses utilize the Internet's capabilities to harvest data about consumers. This process is called web tracking. Over time it has transformed into a very big industry. This study aimed to shed light on how users are tracked, the ways they can protect themselves, and why their data is deemed so significant. There are many ways consumer data can be technologically collected. This includes cookies, browser fingerprinting, and web beacons.

When it comes to types of cookies, the third-party persistent cookie is the most notorious for web tracking. Over time, this data structure can pick up a lot of information. As the underlying data is accessed by many website authors, user "profiles" begin to form. These profiles are used for many different business purposes. To make matters worse, this is typically done without the consumers consent or knowledge. Many countries around the world do not require websites to disclose the use of tracking cookies.

Browser fingerprinting involves gathering information about a user through their browser. Businesses rely on the fact that statistically, it is quite uncommon for two browsers to share the same signature.

As time has gone by, both the defenders and perpetrators of web tracking have had come up with more innovative measures to gain the upper hand. A product of this quest includes the web beacon. Its ability to quickly and stealthily collect and disclose user information has recently made it one of the more popular web tracking methods.

It should be mentioned that the attack surface of web tracking methods are not limited to a website. As computers have become more advanced, embedding interfaces inside emails, SMS messages, and videos have become fairly mainstream. Because similar technologies are often used in these instances, they share the same exploits. For example, web beacons can be embedded inside the body of an email. The same requests are made inside the body of an email as they are inside a web page. Web beacons can provide an advertiser metrics such as the "read" status of the email or if a user clicked on a particular email link. Fortunately, a lot of same defensive measures apply to both contexts.

While cookies, browser fingerprinting, and web beacons are effective tools for consuming data, defenses can be deployed. Third-party persistent cookies can be removed from a browser at any point. There are also popular browser plugins that offer quite the protective measures. Privacy Badger automatically blocks tracking cookies by preventing certain content from loading within a browser [2]. Disconnect detects when a browser requests a resource other than those necessary for the site to load appropriately [2]. Browser fingerprinting's success is based on the fact that most browsers create a unique signature. The more a user's browser fingerprint "blends in" with that of other browsers, the harder it will be for advertisers to collect meaningful data. Surfing the web in "incognito mode" allows consumers to share the same set of browser setting values in a preconfigured manner [2]. Disabling JavaScript will hide valuable information such as active plugins and fonts. Anti-malware can offer an indirect defense by disallowing unwanted processes from adding to the uniqueness of a browser signature. Ghostery can disable web beacons on a single website or across all websites [4]. Google believe it or not, also implements protective measures by default. The company acts as a guard against web beacons collecting information associated with the end consumer.

As the digital age continues to develop, online consumer data will be a highly sought-after product. It is used for better business decision metrics, sold in its raw form, and packaged for malicious purposes, amongst other usages. Once again, the goal of this study is not to forbid users from using the Internet. It is to enlighten them on how their information is tracked and the ways they can surf the web in a more private manner.

7.References:

[1] Anon, D., 2018. How Cookies Track You Around The Web & How To Stop Them | Privacy.Net. [online] Privacy.net. Available at: <https://privacy.net/stop-cookies-tracking> [Accessed 15 November 2020].

[2] Pixel Privacy. 2020. Browser Fingerprinting: What Is It And What Should You Do About It?. [online] Available at: <https://pixelprivacy.com/resources/browser-fingerprinting/> [Accessed 15 November 2020].

[3] Frew, J., 2016. How Advertisers Use Web Beacons To Track You On The Web And In Emails. [online] MakeUseOf. Available at: <https://www.makeuseof.com/tag/how-web-beacons-track-web/> [Accessed 15 November 2020].

[4] Gralla, P., 2015. How To Stop Ad Trackers And Beacons Dead In Their Tracks. [online] Computerworld. Available at: <https://www.computerworld.com/article/2974461/how-to-stop-ad-trackers-and-beacons-dead-in-their-tracks.html> [Accessed 15 November 2020].

[4] En.wikipedia.org. 2020. Dick's Sporting Goods. [online] Available at: <https://en.wikipedia.org/wiki/Dick%27s_Sporting_Goods> [Accessed 15 November 2020].

[5] Alexa.com. 2020. Dickssportinggoods.Com Competitive Analysis, Marketing Mix And Traffic - Alexa. [online] Available at: <https://www.alexa.com/siteinfo/dickssportinggoods.com> [Accessed 15 November 2020].

[6] Import.io. 2015. What Is Data, And Why Is It Important? | Import.Io. [online] Available at: <https://www.import.io/post/what-is-data-and-why-is-it-important/> [Accessed 15 November 2020].

[7] "Billandpauls.Com Competitive Analysis, Marketing Mix And Traffic - Alexa". *Alexa.Com*, 2020, https://www.alexa.com/siteinfo/billandpauls.com.

[8] Agrawal, A., 2016. Why Data Is Important For Companies And Why Innovation Is On The Way. [online] Inc.com. Available at: <https://www.inc.com/aj-agrawal/why-data-is-important-for-companies-and-why-innovation-is-on-the-way.html> [Accessed 15 November 2020].

[9] Zawadziński, M., 2015. The Truth About Online Privacy: How Your Data Is Collected, Shared, And Sold - Clearcode Blog. [online] Clearcode | Custom AdTech and MarTech Development. Available at: <https://clearcode.cc/blog/online-privacy-user-data/> [Accessed 15 November 2020].

[10] Mentalfloss.com. 2020. 11 Ways Your Online Data Is Being Used Right Now. [online] Available at: <https://www.mentalfloss.com/article/84790/11-ways-your-online-data-being-used-right-now> [Accessed 15 November 2020].

[11] The Tor Project | Privacy & Freedom Online. 2020. Available at: < https://www.torproject.org/> [Retrieved 28 November 2020].