



The State of Web Privacy

Threats & Defenses

Umar Iqbal
University of Washington

 @umaarr6

Targeted Ads



Running shoes



<https://www.shoecarnival.com> › athletics_and_sneakers

Men's Running Shoes - Shoe Carnival

Shop **running shoes** for men at Shoe Carnival. Shop great deals on men's **running shoes**, in store & online. Members get FREE shipping!

<https://www.colehaan.com> › Men › Men's Performance ▾

Men's Running, Training, & All-Terrain Shoes | Cole Haan

Men's **running shoes** combine everything there is to love about undoubtedly stylish running-wear and comfort whether you are going long distance routes or ...

<https://www.t3.com> › Features › Fitness

Best running shoes 2022 from Adidas, Nike, ASICS and more

Best **running shoes** 2022 with running trainers from Adidas, Nike, ASICS and more. The best **running shoes** for jogging, training and racing reviewed and ranked by ...

<https://www.famousfootwear.com> › shoes › womens › run...

Women's Running Shoes - Famous Footwear

Products 1 - 44 of 322 — Shop **Running Shoes** for Women. Discover the latest styles from your favorite brands at Famous Footwear. Free Shipping available for ...

<https://runkeeper.com> › cms › health › why-you-should... ▾

Why You Should Rotate Your Running Shoes - Runkeeper

Why? Rotating your **running shoes** lowers your injury risk, makes your shoes last longer and optimizes your performance in each run. Alternating shoes reduces ...



techcrunch.com

Join TechCru

foxnews.com

U.S. Politics Media Opinion Business Entertainment Sports Lifestyle

Hot Topics HUNTER BIDEN ROAD TO 2024 FTX COLLAPSE WORLD CUP

MS

TUESDAY STEAL

FRESH FOAM VONGO 4
(Indigo and Sedona)

\$59.88
~~\$139.99~~

SHOP WOMEN ▶ SHOP MEN ▶ SHOP KIDS ▶

Agenda

How are advertisers able to infer your interests?

What technical mechanisms do advertisers use to get that information?

How prevalent are these practices on the internet?

How can users protect their privacy on the internet?

Third-party tracking

Websites load 100s of third-party organizations

Cross-site tracking

Infer sensitive information

Medical history

Financial situation

Weight conscious

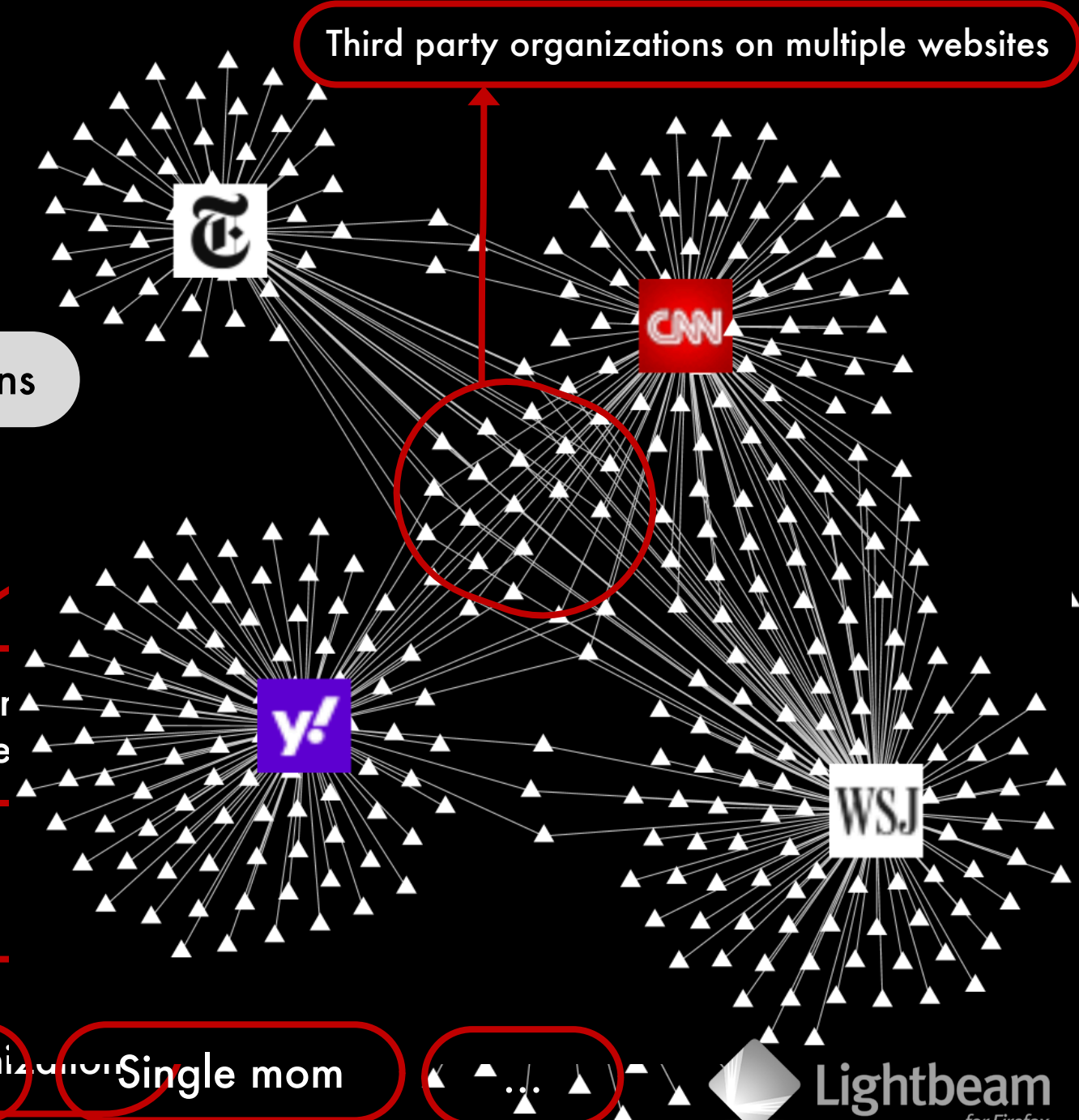
Gambler

Religious cause donor

Drug addict

Single mom

Third party organizations on multiple websites



Third party orgar multiple we

Third party organizations

User data is sold online!

Online Behavior Advertising

"I Lost 12 Lbs in 2 Weeks"
Read How I Went from Flab to Fab

LOCAL SINGLES ALERT
Meet Attractive Single People that Live Near You!

3 MINUTE
Start your application!

1 APPLY NOW Completely Online
2 GET APPROVED Receive Fast Approval
3 GET FUNDS In Less Than an Hour!

Privacy & Security PROTECTED

\$120+ billion yearly revenue [IAB]

Data Brokers

DMDATABASES.COM
Your ultimate source for marketing databases

TOP SELLING AILMENT EMAIL LISTS / MAILING LISTS

Wheelchair Users Mailing List: over 1,000,000
Insulin Users Mailing List: over 1,000,000
Back Pain Sufferers Mailing List: over 1,000,000
Sleep Apnea Sufferers Mailing List: over 1,000,000
Asthma Sufferers Mailing List: over 1,000,000
Asthma Medication Users Mailing List: over 1,000,000
Rheumatoid Arthritis Sufferers Mailing List: over 700,000
Diabetics Mailing List: over 12,000,000
Bladder Control Sufferers Mailing List: over 340,000
ADHD Ailment Mailing List: over 500,000
Clinical Depression Sufferers Mailing List: over 1,800,000

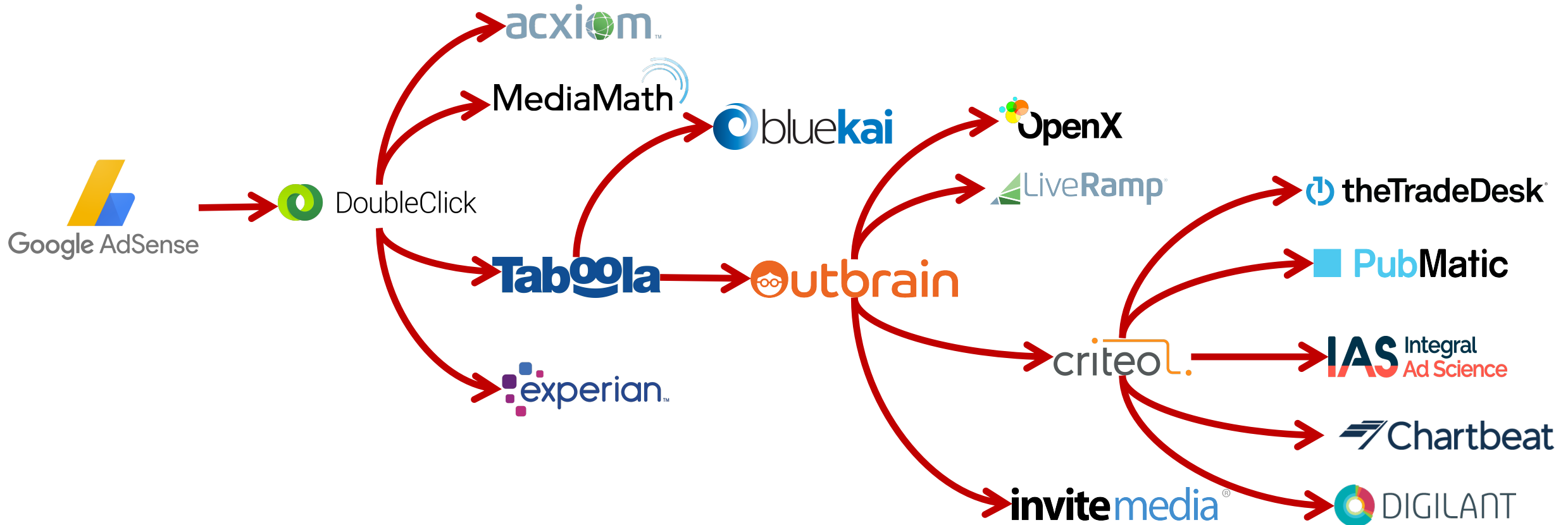
Sold for less than \$1K

The modular web

Add functionality by embedding third-party organizations



Advertising & tracking services



"Trusted" organizations



DoubleClick



Criteo



Taboola

invite media



DoubleVerify



AppNexus



JWPLAYER

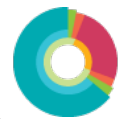


PubMatic



Quantcast

Google AdSense



DIGILANT

Google Analytics

mopub

OpenX

acxiom

bluekai



Akamai

Outbrain

MediaMath



jQuery
write less. do more.

bluekai



Akamai

Outbrain



Bing ads



Chartbeat



Bootstrap

adform



theTradeDesk



Chartbeat



Google Tag Manager



amazon advertising

DISQUS

Google Tag Manager



Google AdMob

LiveRamp

experian

choozle

IAS

Integral

Ad Science

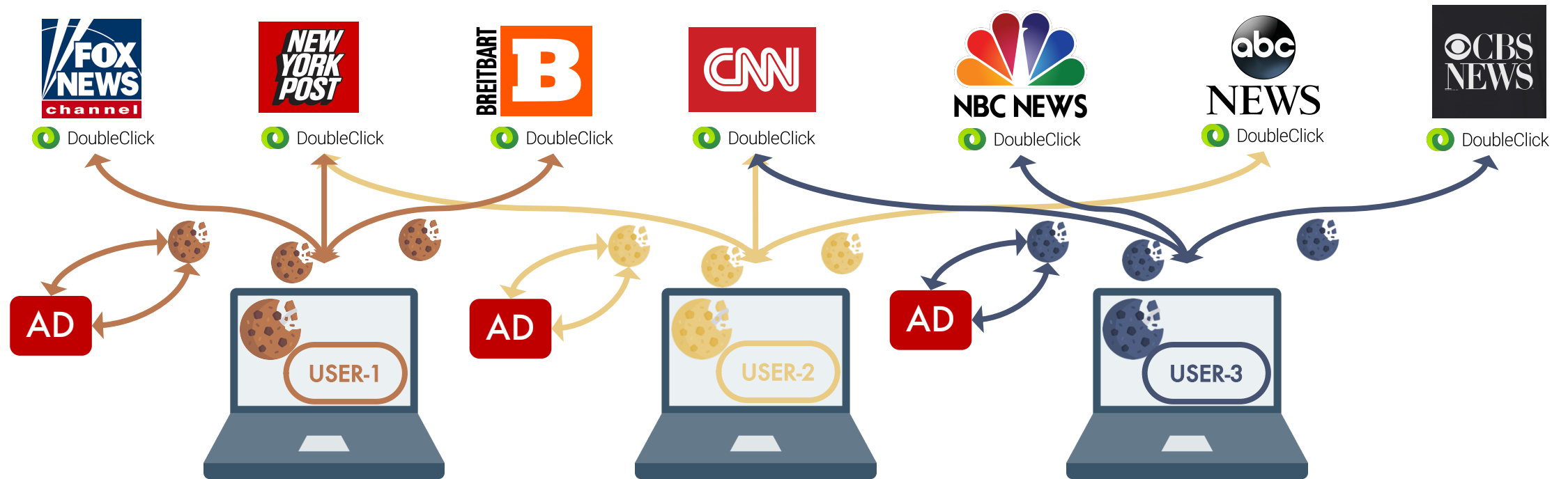
INMOBI



YouTube

Tracking mechanisms

Online tracking



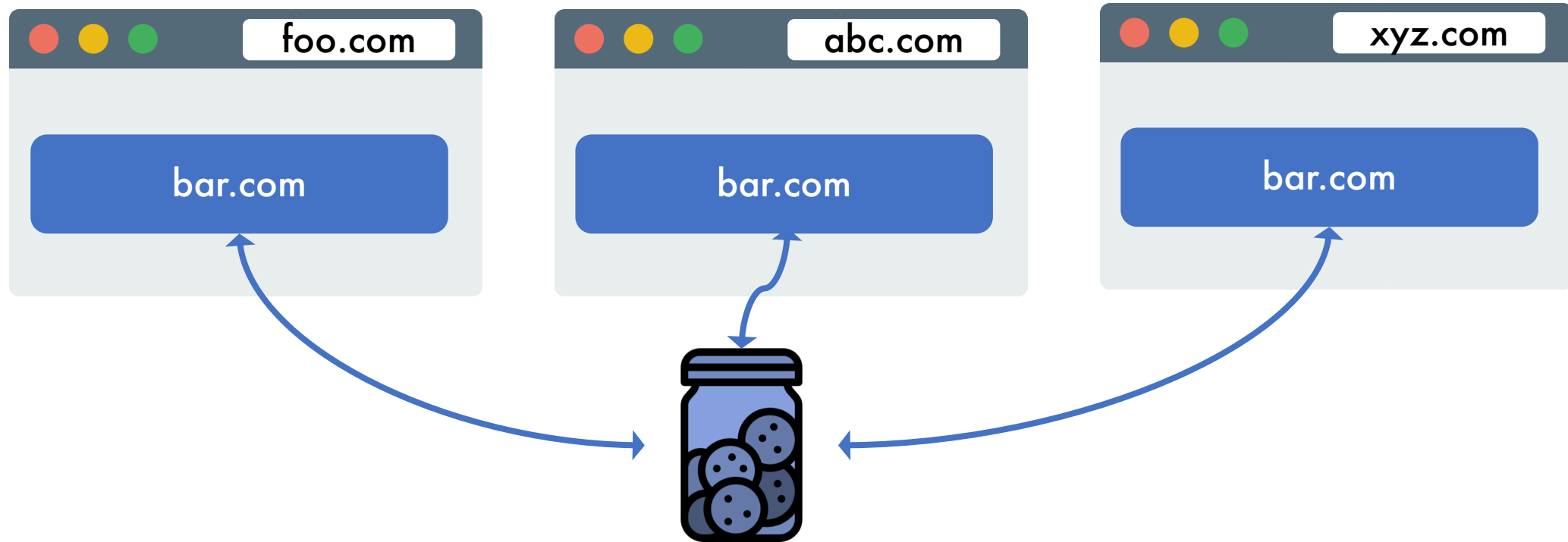
 DoubleClick Database

USER-1	USER-2	USER-3
FoxNews	NewYorkPost	CNN
NewYorkPost	CNN	NBC News
Brietbart	ABC News	CBS News

Cookies

Store state in the browser
Shopping carts, User preferences, Online tracking

Third-party cookies: belong to the domain of embedded content (image, iframes)



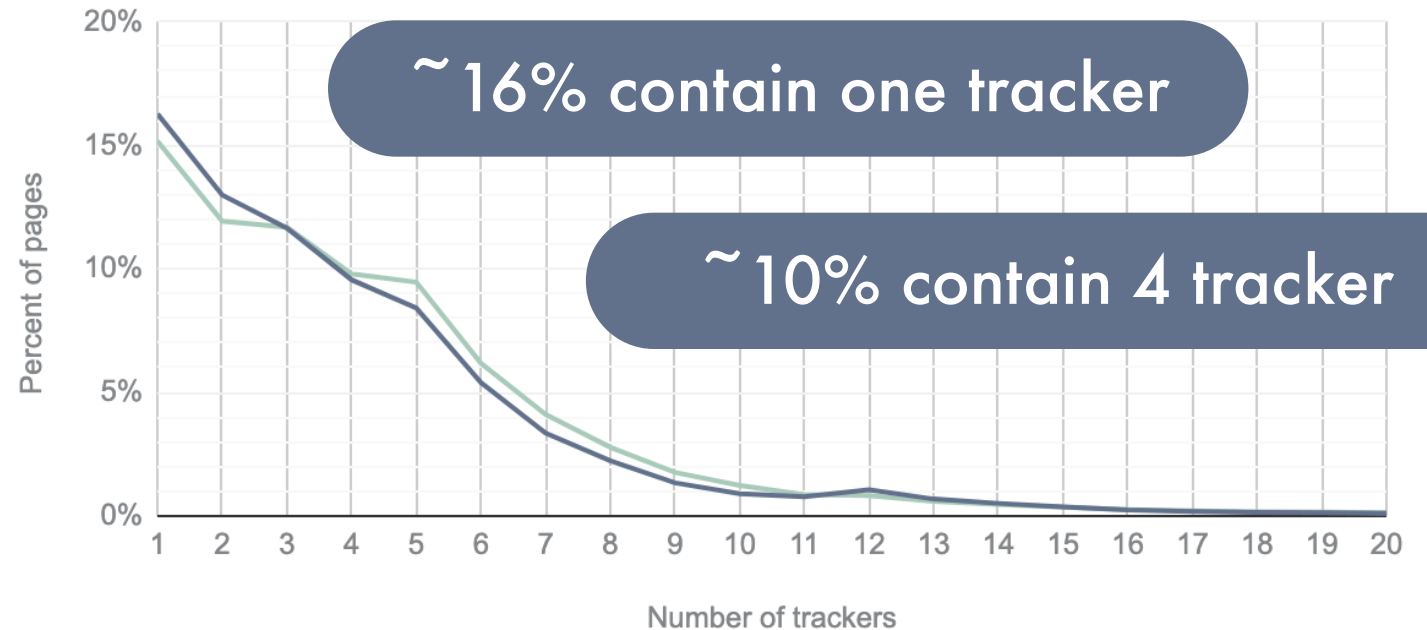
Tracking prevalence^[1]

82% of all websites contain at least one tracker

Number of trackers per website

Web Almanac 2022: Privacy

desktop mobile



Long tail with more than 10 trackers

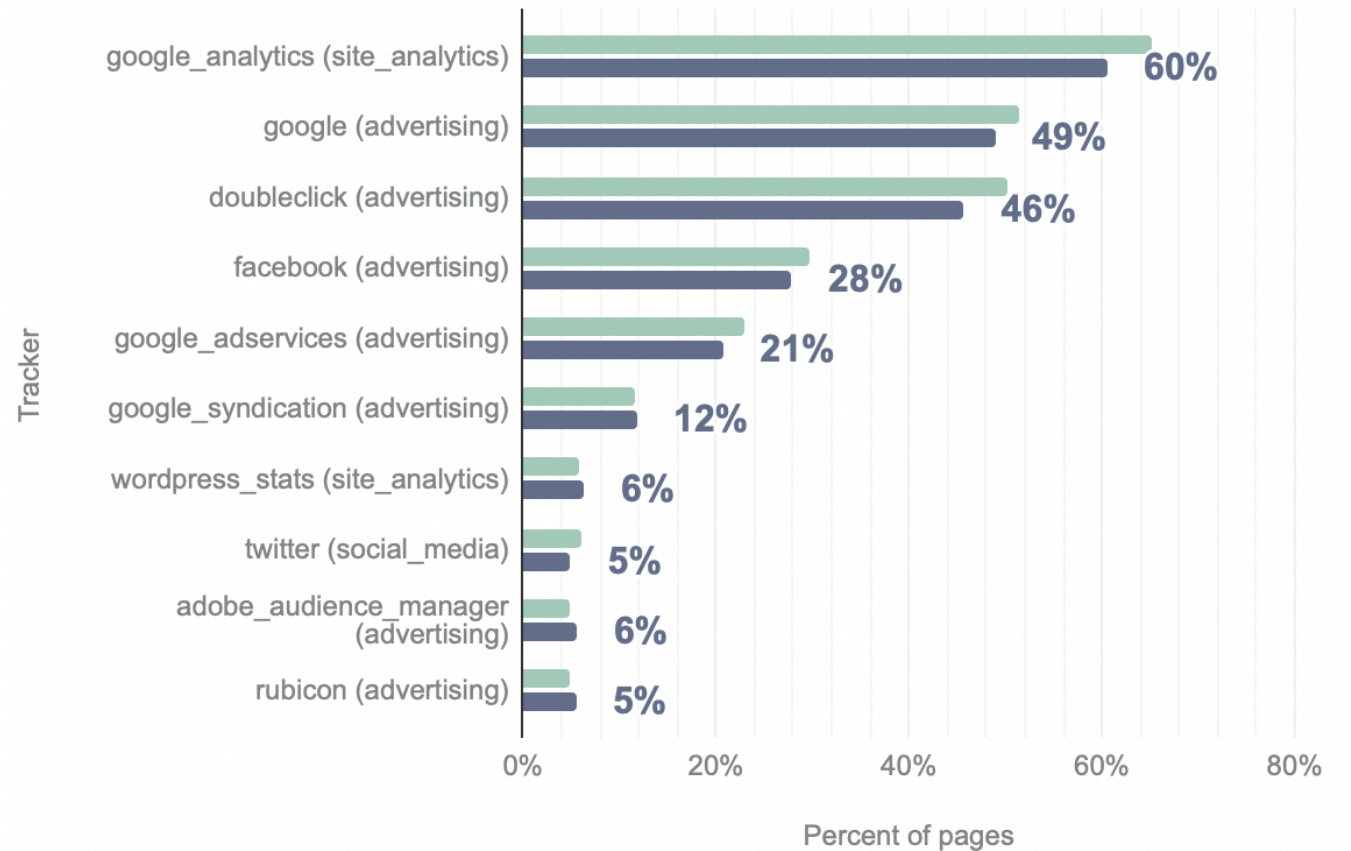
[1] Measurements conducted on ~8 Million websites: <https://almanac.httparchive.org/en/2022/privacy>

Tracking prevalence^[1]

Most common trackers

Web Almanac 2022: Privacy

■ desktop ■ mobile

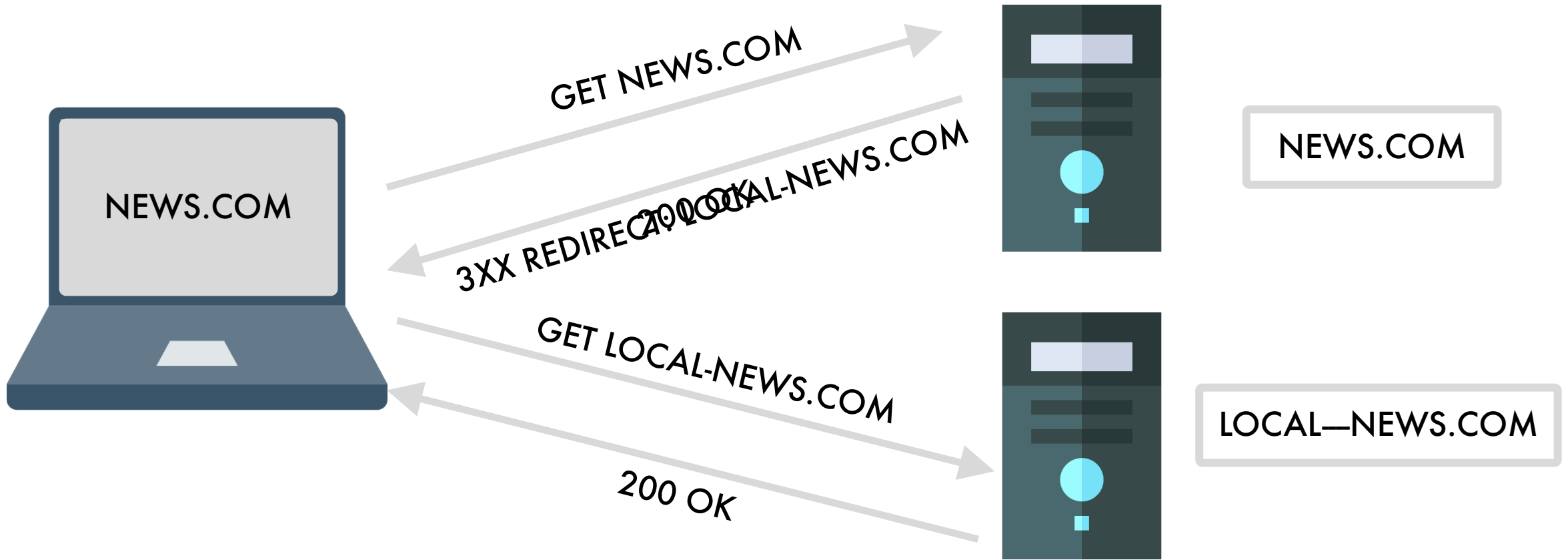


[1] Measurements conducted on ~8 Million websites: <https://almanac.httparchive.org/en/2022/privacy>

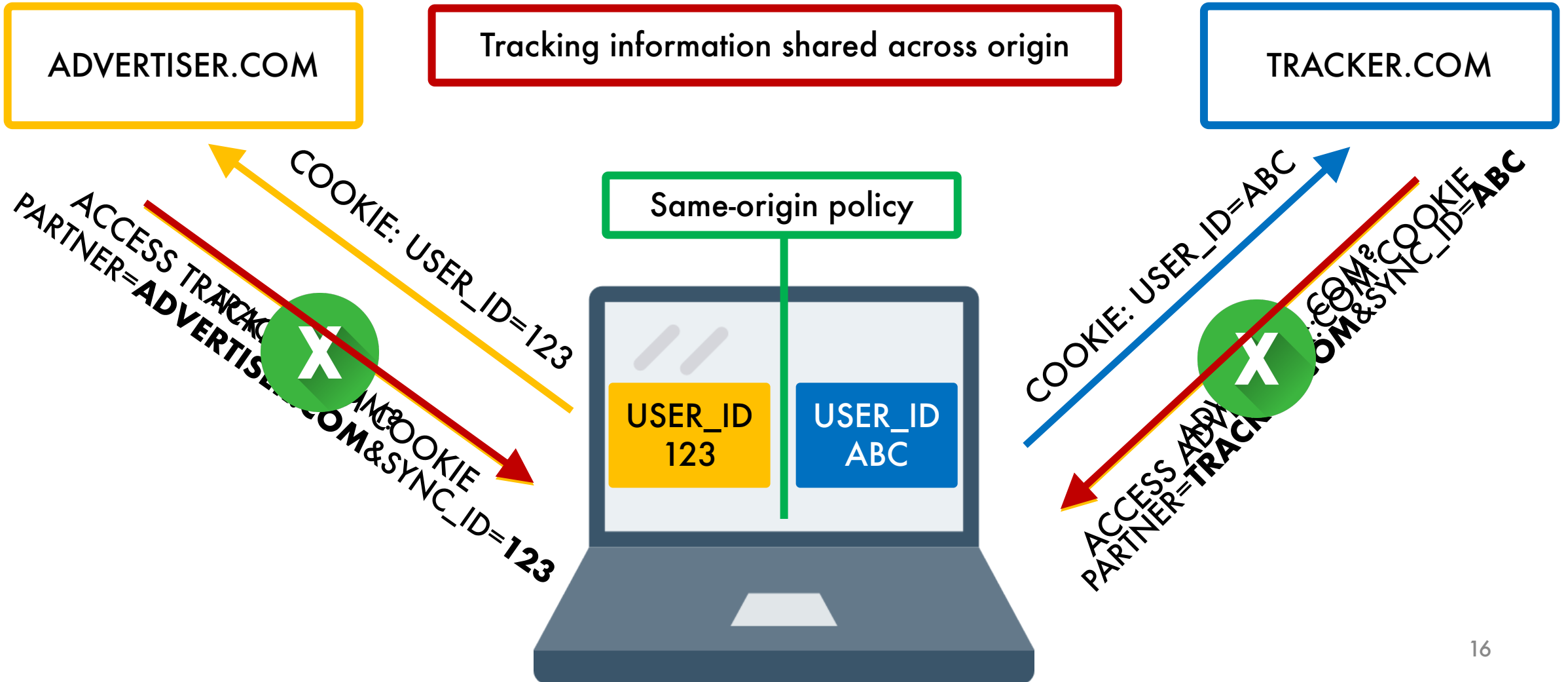
Trackers share data with each other

Even with trackers not present on the website to enhance their
coverage!

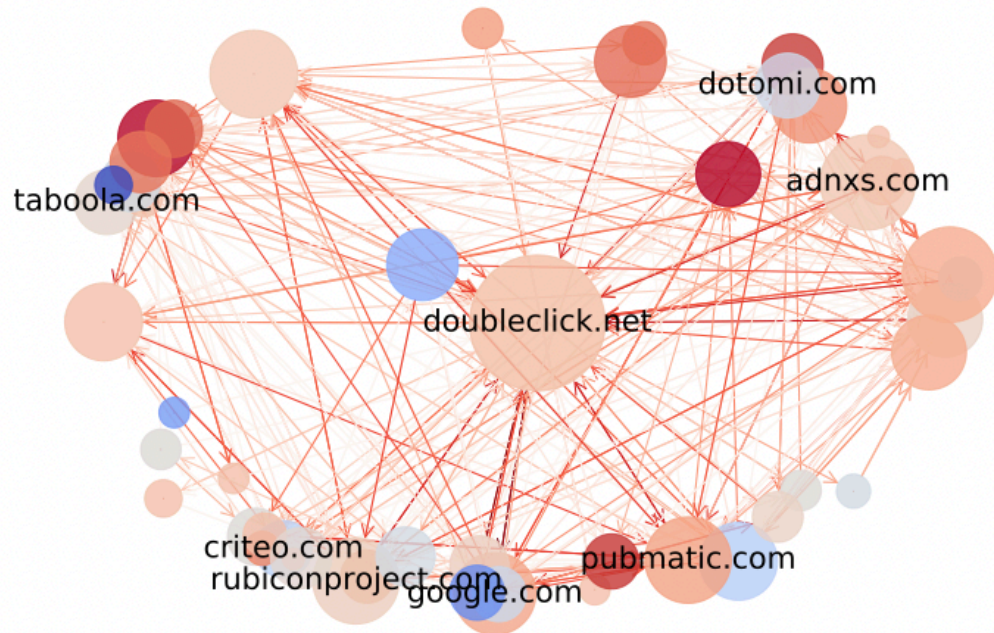
What are request chains?



Cookie syncing



Tracking prevalence^{[1][2]}

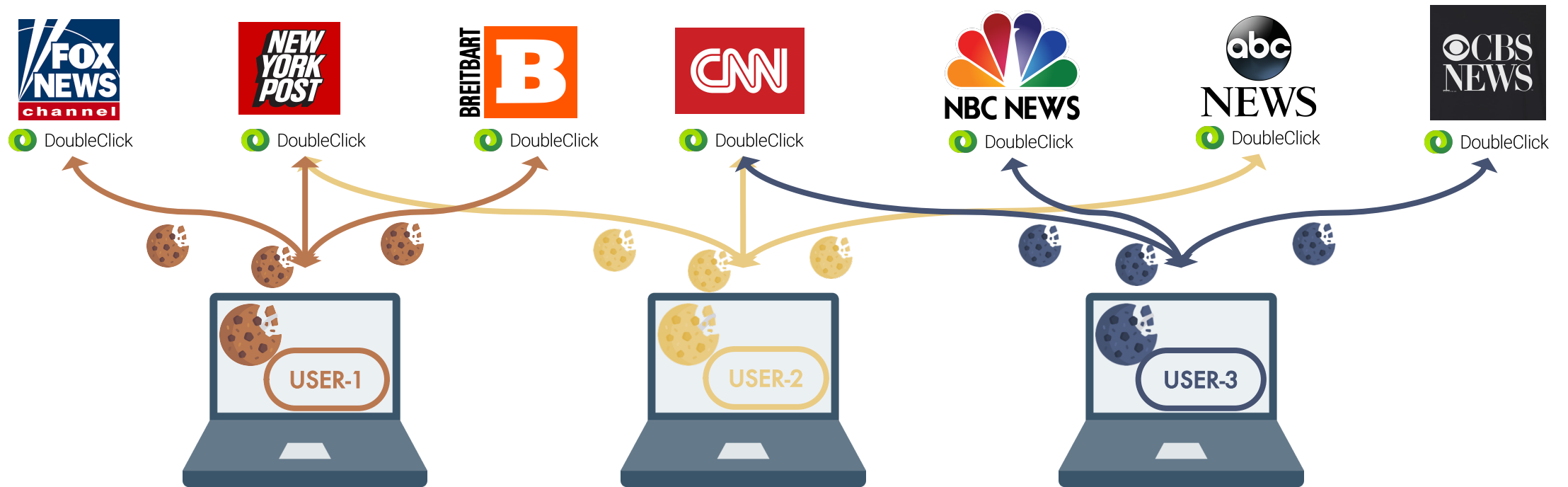


doubleclick.net synced
its cookies with
118+ third parties
(2016)

Privacy protections by main-stream browsers

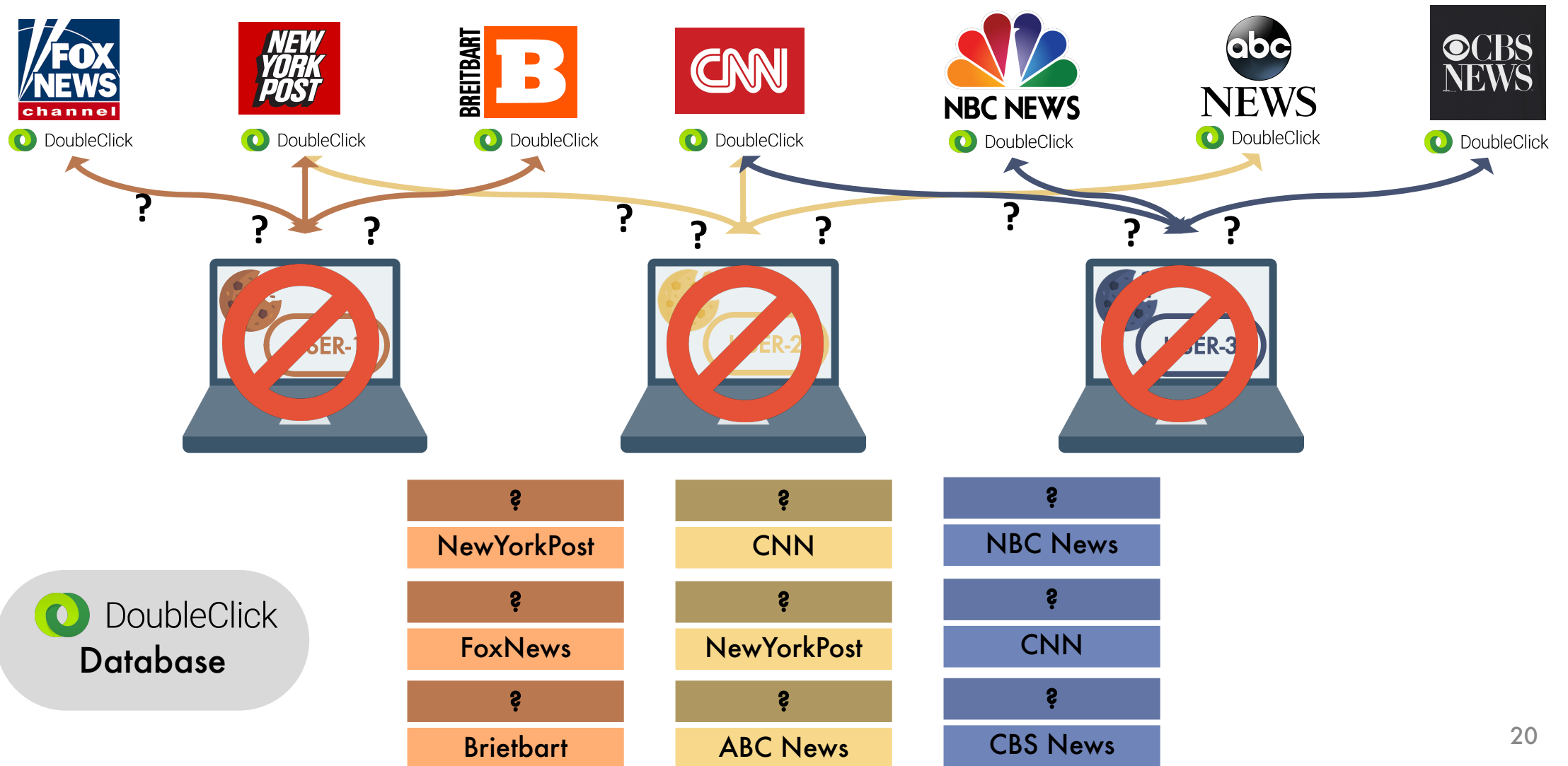
How can users protect their privacy?

Cookie-based tracking



USER-1	USER-2	USER-3
FoxNews	NewYorkPost	CNN
NewYorkPost	CNN	NBC News
Brietbart	ABC News	CBS News

Impact of cookie blocking



Cookie Blocking



Full Third-Party Cookie Blocking and More

Mar 24, 2020 by John Wilander [@johnwilander](#)



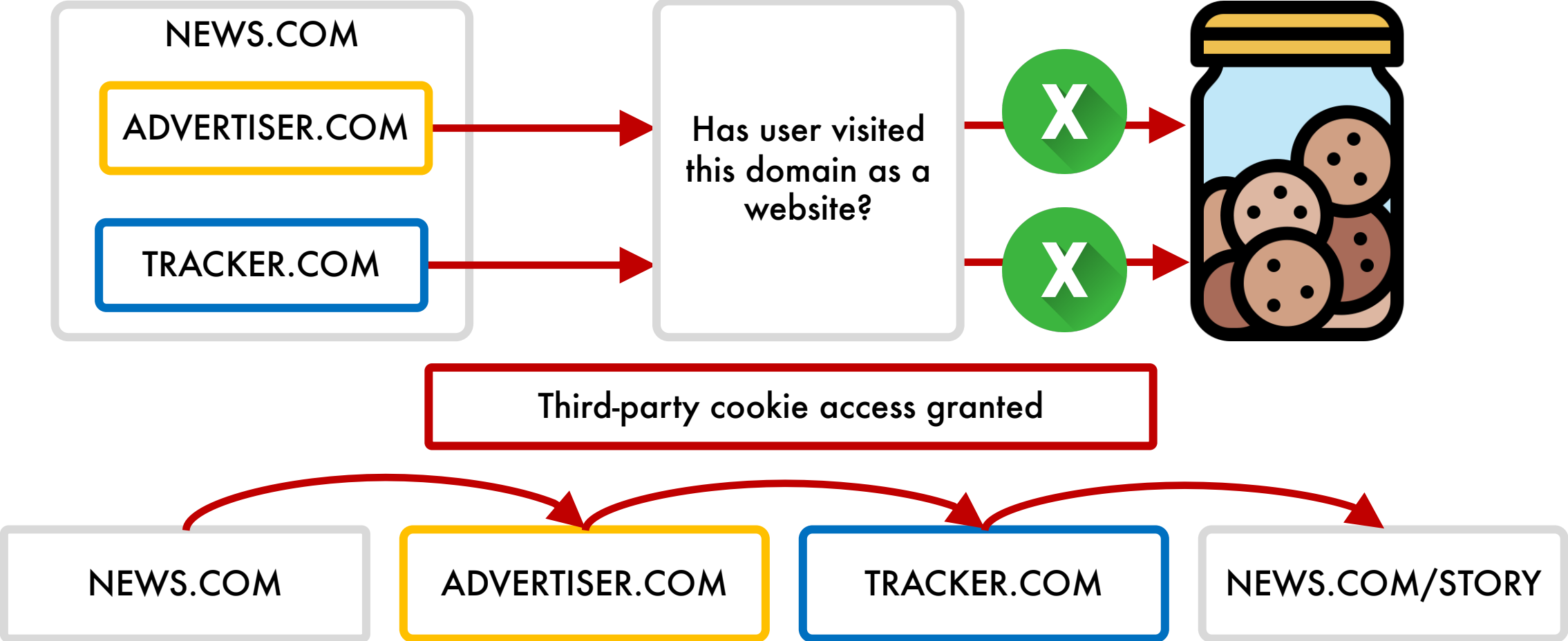
Today's Firefox Blocks Third-Party Tracking Cookies September 3, 2019



Building a more private web: A path towards making third party cookies obsolete

Tuesday, January 14, 2020

Bounce tracking: bypassing cookie blocking



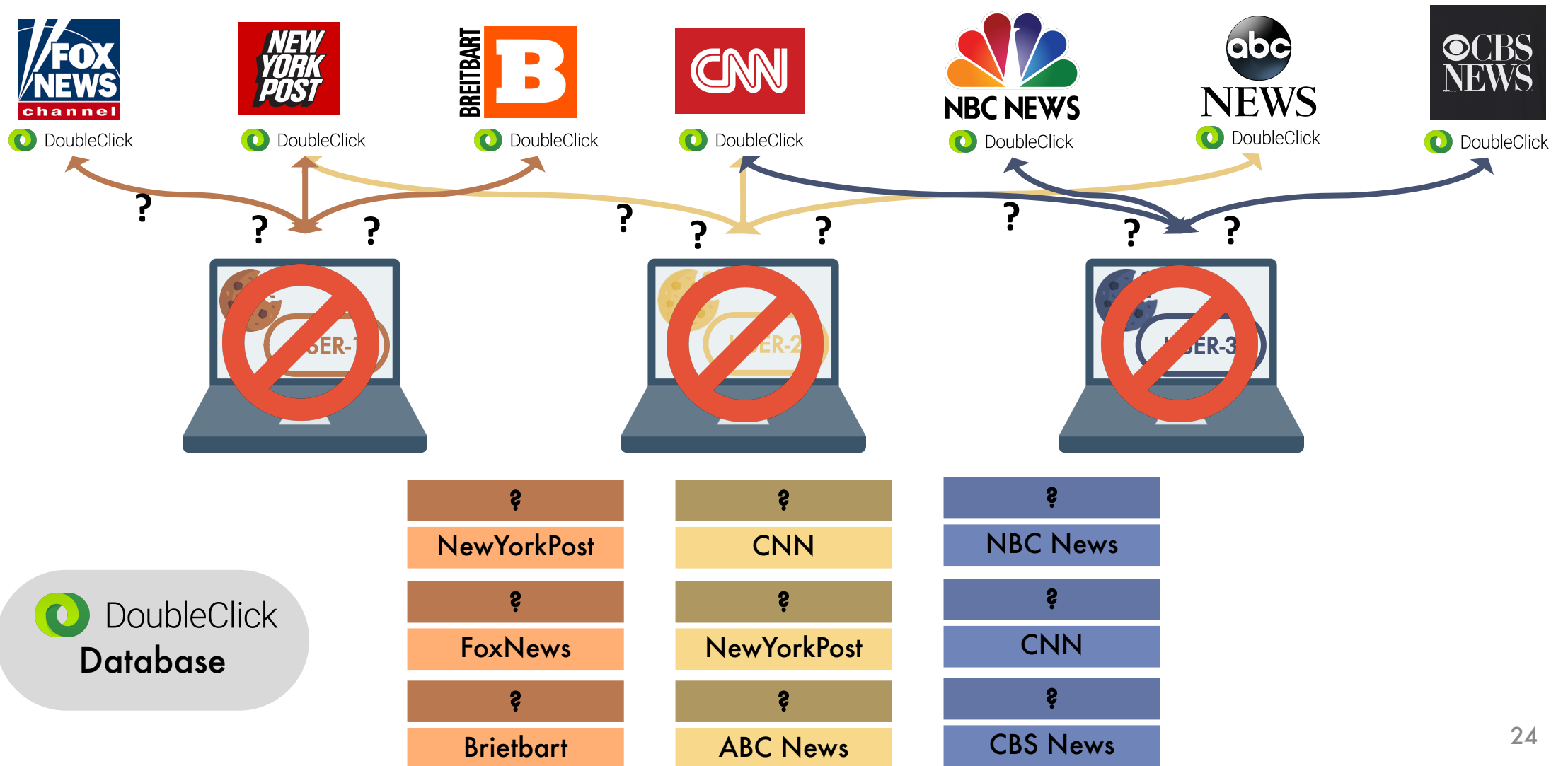
Bounce tracking prevalence^[1]

~ 14 vendors involved
in this practice on top
10K websites

Top Domains	Spoofed Safari # of websites
googleadservices.com	3073
adsrvr.org	1377
adform.net	322
flashtalking.com	141
queue-it.net	9
ojrq.net	-
bngpt.com (NSFW)	2

[1] <https://umariqbal.com/papers/khaleesi-usenix2022.pdf>

Impact of cookie blocking



Browser fingerprinting: alternative tracking techniques

<https://amiunique.org>

Attribute	Example
User agent	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/64.0.3282.119 Safari/537.36
List of plugins	Plugin 1: Chrome PDF Plugin. Plugin 2: Chrome PDF Viewer. Plugin 3: Native Client. Plugin 4: Shockwave Flash...
Cookies enabled	yes
Use of local/session storage	yes
Timezone	-60 (UTC+1)
Screen resolution and color depth	1920x1200x24
List of fonts	Abyssinica SIL,Aharoni CLM,AR PL UMing CN,AR PL UMing HK,AR PL UMing TW...
Platform	Linux x86_64
Do Not Track	yes
Canvas	Cwm fjordbank glyphs vext quiz, ☺ Cwm fjordbank glyphs vext quiz, ☺
WebGL Vendor	NVIDIA Corporation
WebGL Renderer	GeForce GTX 650 Ti/PCIe/SSE2
Use of an ad blocker	yes



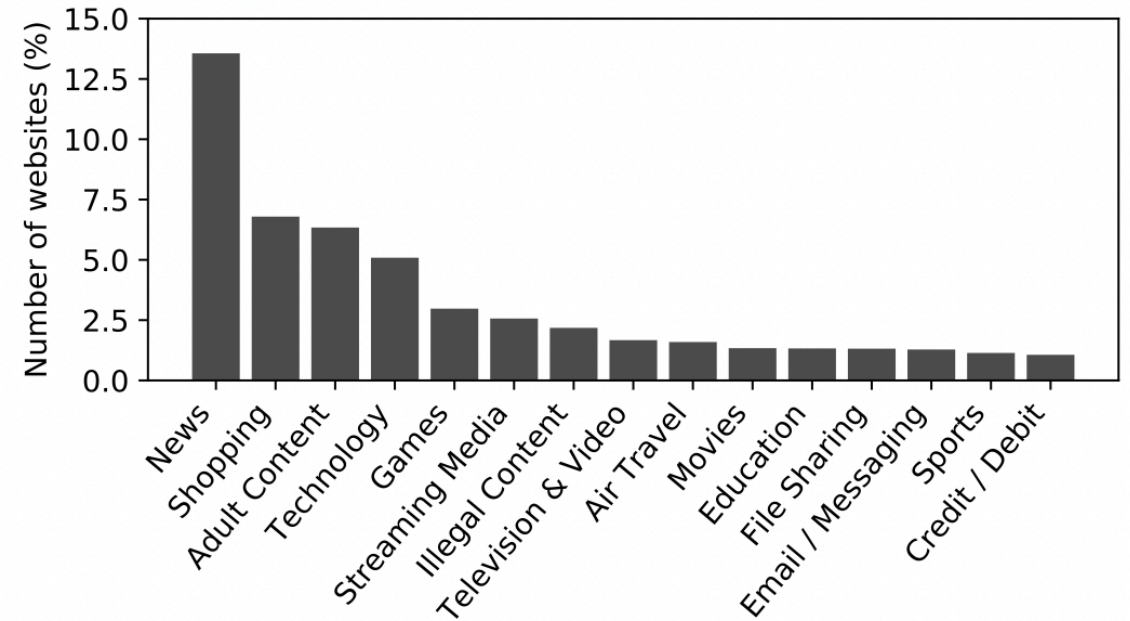
 DoubleClick Database

USER-1	USER-2	USER-3
FoxNews	NewYorkPost	CNN
NewYorkPost	CNN	NBC News
Brietbart	ABC News	CBS News

Browser fingerprinting prevalence^[1]

>1/3rd of popular websites deploy fingerprinting scripts

Rank Interval	Websites (count)	Websites (%)
1 to 1K	266	30.60%
1K to 10K	2,010	24.45%
10K to 20K	981	11.10%
20K to 50K	2,378	8.92%
50K to 100K	3,405	7.70%
1 to 100K	9,040	10.18%



[1] <https://umariqbal.com/papers/fpinspector-sp2021.pdf>

Browser fingerprinting prevalence^[1]^[2]

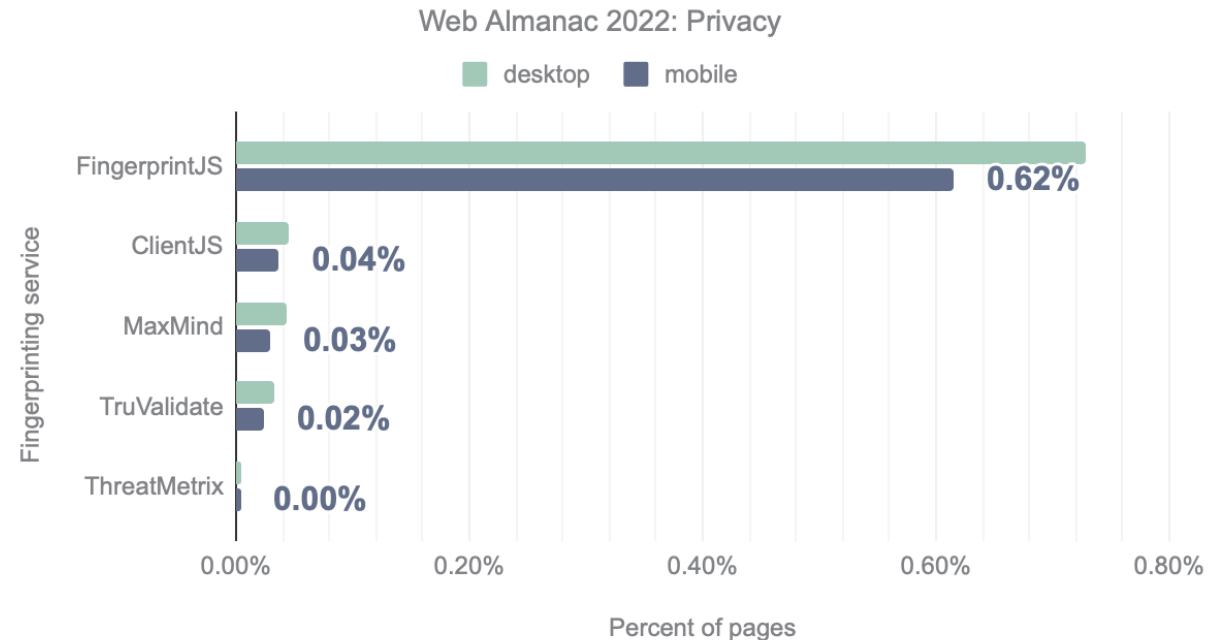
Vendor Domain	Tracker	Websites (count)
doubleverify.com	Y	2,130
adsafeprotected.com	Y	1,363
alicdn.com	N	523
adsco.re	N	395
yimg.com	Y	246
2,344 others	Y(86)	5,702
Total		10,359 (9,040 distinct)

Popular trackers & Ad fraud services

DoubleVerify

Integral Ad Service

Fingerprinting services usage



[1] <https://umariqbal.com/papers/fpinspector-sp2021.pdf>

[2] <https://almanac.httparchive.org/en/2022/privacy>

Privacy protections by browser extensions

How can users protect their privacy?

Privacy-enhancing tools



Browser extensions
&
Privacy focused browsers

EasyList
DISCONNECT.

Filter lists of known
advertising and tracking
domains

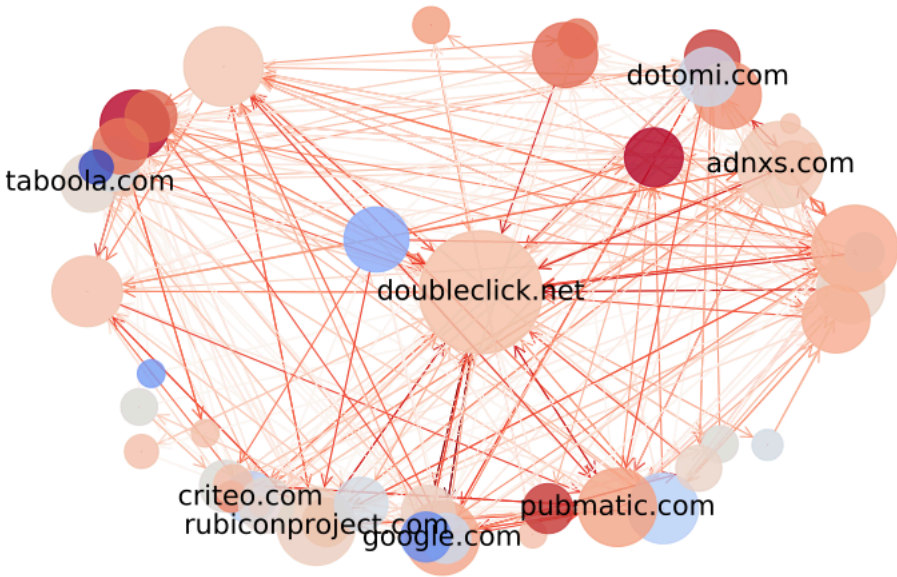
ⓘ daijiworld.com
#6825 opened 4 days ago by ghajini
ⓘ indiatimes.com
#6822 opened 4 days ago by ghajini
ⓘ bikeroar.com
#6790 opened 6 days ago by DandelionSprout

Crowdsourced

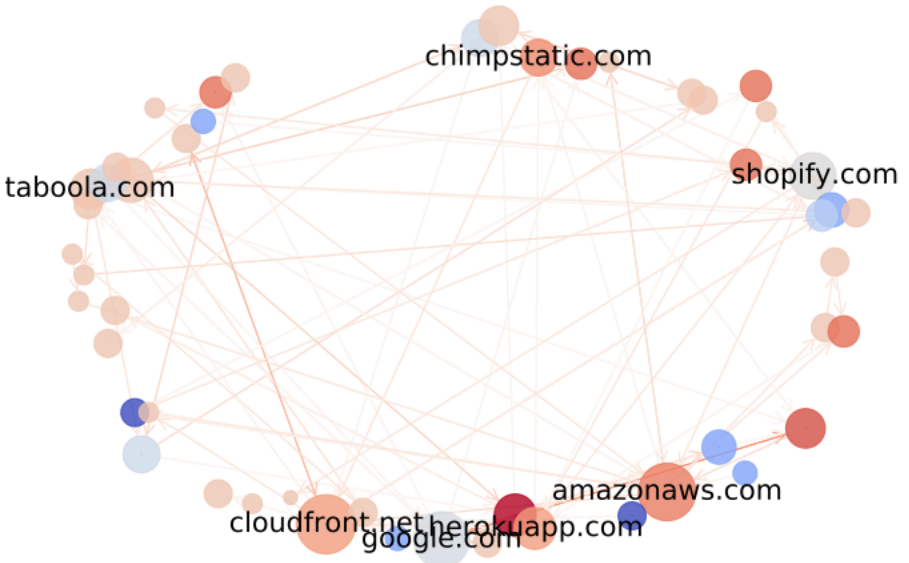
[canvas fingerprinting]
canvas width, height > 15 px
calls to save, restore = None
[Blocked]

Heuristic based

Tracker blocking



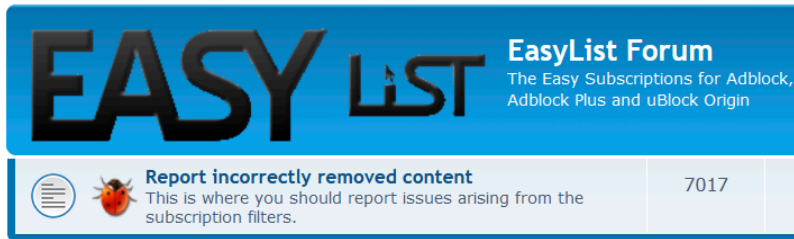
(a) No Blocking



(b) Filter Lists

Limitations of ad/tracker blockers [1]

Brittle



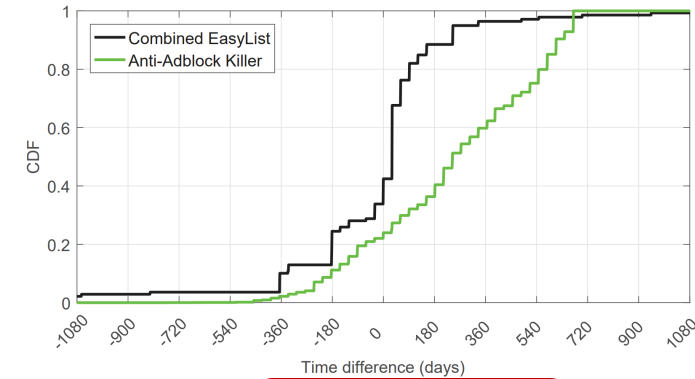
Mistakes

```
||1juxicz.tkepyc.com, domain=pch.com
||la112zu.tkepyc.com, domain=pch.com

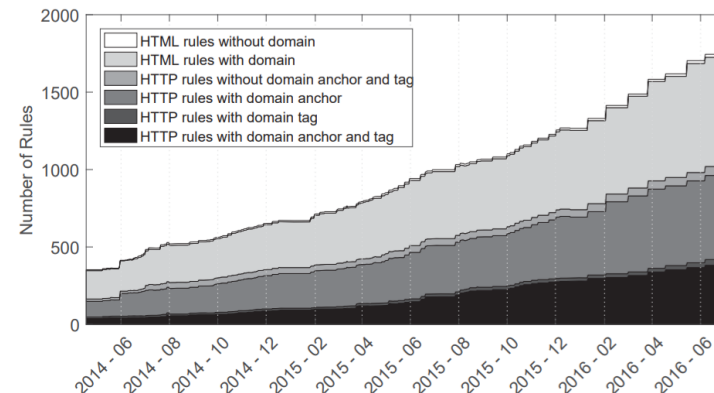
var n = "UNICODE STRING",
i = e.getContext("2d");
i.save(), i.rect(0,0,10,10), i.rect(2,2,6,6),
t.push(!1 === i.isPointInPath(5, 5, "evenodd")
? "yes" : "no"), i.restore(), i.save();
var r = i.createLinearGradient(0, 0, 200, 0);
.....
i.shadowColor="rgb(85,85,85)", i.shadowBlur=3,
i.arc(500,15,10,0,2*Math.PI,!0), i.stroke(),
i.closePath(), i.restore(), t.push(e.toDataURL())
```

Circumvention

Manual




Slow



Redundant

[1] <https://umariqbal.com/papers/adbloc-imc2017.pdf>

Why is it challenging to protect?




Adblock Detected

We have detected that you are using an a plugin to disable advertising from loading

The revenue earned from advertising enables us to pro you're trying to reach on this website. In order to view that you disable adblock in plugin se

I HAVE DISABLED ADBLOCK FOR TH



Adblock Detected

[Forbes Login](#)

Yes, I'll Whitelist

or

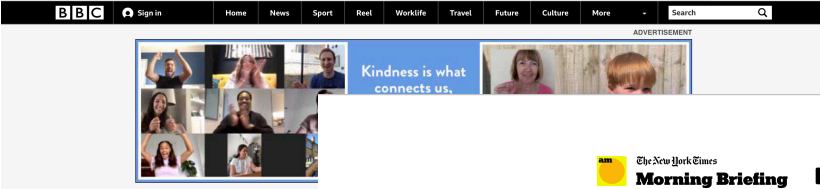
[Log in](#) [Google](#)

We will never post on your behalf without first obtaining your permission. By joining through Facebook or Google sign-up, you agree to the [Terms Of Membership](#) for all Forbes Media sites.


BLEEPINGCOMPUTER

Ad Network Uses DGA Algorithm to Bypass Ad Blockers and Deploy In-Browser Miners


Adversarial Evasions



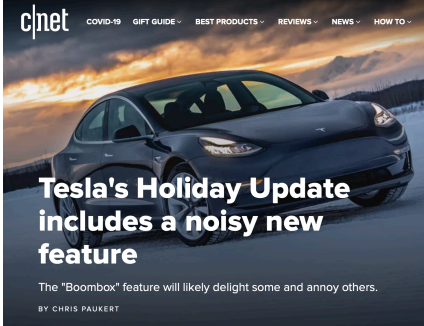
Kindness is what connects us.



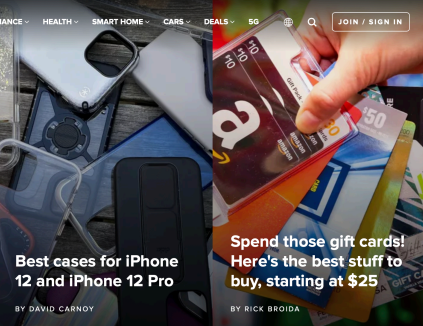
The New York Times Morning Briefing [SIGN UP NOW](#)



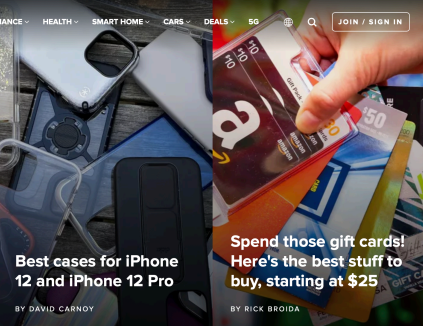
Best Headphones for 2020 [LEARN MORE](#)



Tesla's Holiday Update includes a noisy new feature



Best cases for iPhone 12 and iPhone 12 Pro



Spend those gift cards! Here's the best stuff to buy, starting at \$25

Scalability

How can we address this problem?

Flawed approach

Cannot possibly put all the trackers in the world in a list!

Fundamentally change in approach

May be detect tracking behavior instead of trackers?

Key insight



Execution traces as a signature of tracking behaviors

Interaction with different elements on a webpage
External network requests initiated by a script

.....

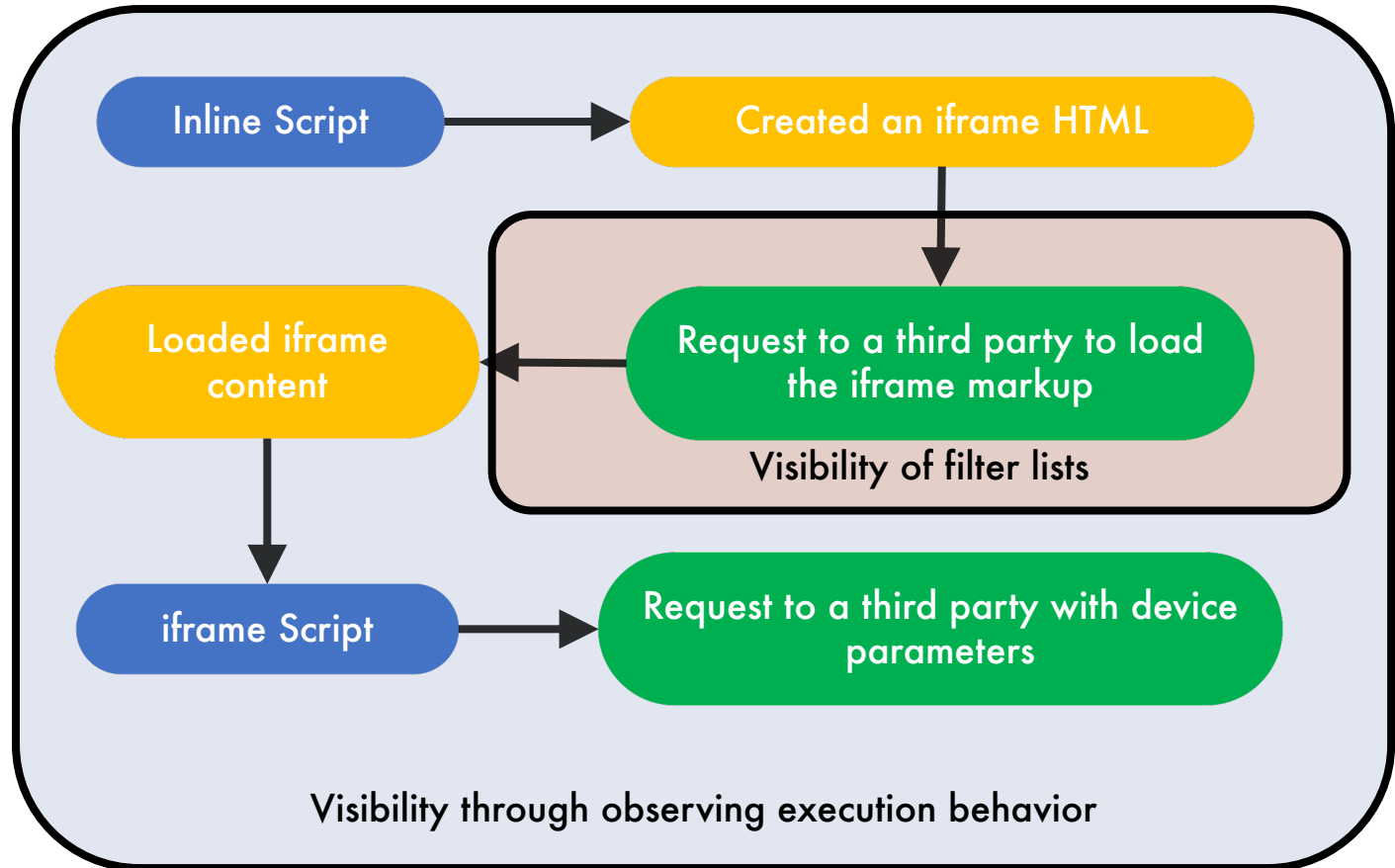
Key insight

```
<script>
  iframe = document.createElement("iframe");
  iframe.src = "tracker.com/load_tracker.html";
  document.body.appendChild(iframe)
</script>

<iframe src="tracker.com/load_tracker.html">
  <script>
    // Canvas font fingerprinting script.
    Fonts = ["monospace", ... , "sans-serif"];
    CanvasElem = document.createElement("canvas");
    CanvasElem.width = "100";
    CanvasElem.height = "100";
    context = CanvasElem.getContext('2d');
    FPDict = {};

    for (i = 0; i < Fonts.length; i++) {
      CanvasElem.font = Fonts[i];
      FPDict[Fonts[i]] = context.measureText("example").width;
    }

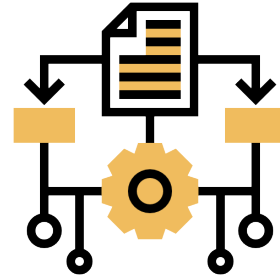
    var img = document.createElement("img");
    img.src = "tracker.com/track_user/?userId={FPDict}";
  </script>
</iframe>
```



Browser instrumentation & Machine learning



Instrument browsers



Machine learning based detection

Crucial for robustness

Capture execution provenance at multiple layers

Reveal any evasion attempts

Crucial for scalability

Detect tracking behavior

Generalize behaviors across trackers

Regulations



Ask for user permission before collecting their data

Regulations

Manage Consent Preferences

+ Strictly Necessary Cookies Always Active

+ Analytics Cookies

+ Functional Cookies

+ Targeting Cookies

Reject All

Consent **Details** **About**

Necessary 22

Necessary cookies help make a website usable by enabling basic functions like page navigation and access to secure areas of the website. The website cannot function properly without these cookies.

Preferences 13

Preference cookies enable a website to remember information that changes the way the website behaves or looks, like your preferred language or the region that you are in.

Statistics 18

Statistic cookies help website owners to understand how visitors interact with websites by collecting and reporting information anonymously.

Marketing 21

Marketing cookies are used to track visitors across websites. The intention is to display ads that are relevant and engaging for the individual user and thereby more valuable for publishers and third party advertisers.

Deny **Allow selection** **Allow all**

Powered by Cookiebot by Usercentrics



Data Rights Protocol

Standardizing and streamlining consumer data rights requests



The State of Web Privacy

Threats & Defenses

Umar Iqbal

 umariqbal.com

 @umaarr6