Leaky Forms: A Study of Email and Password Exfiltration Before Form Submission

Asuman Senol
imec-COSIC, KU Leuven
asuman.senol@esat.kuleuven.be

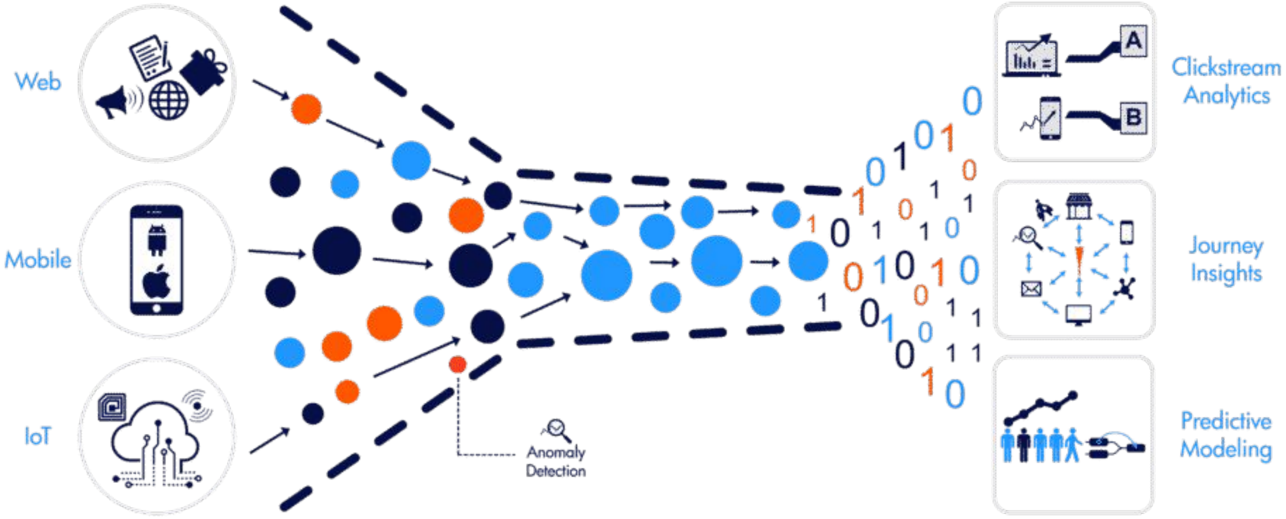
Gunes Acar
Radboud University
g.acar@cs.ru.nl

Mathias Humbert
University of Lausanne
mathias.humbert@unil.ch

Frederik Zuiderveen Borgesius
Radboud University
frederikzb@cs.ru.nl

Background

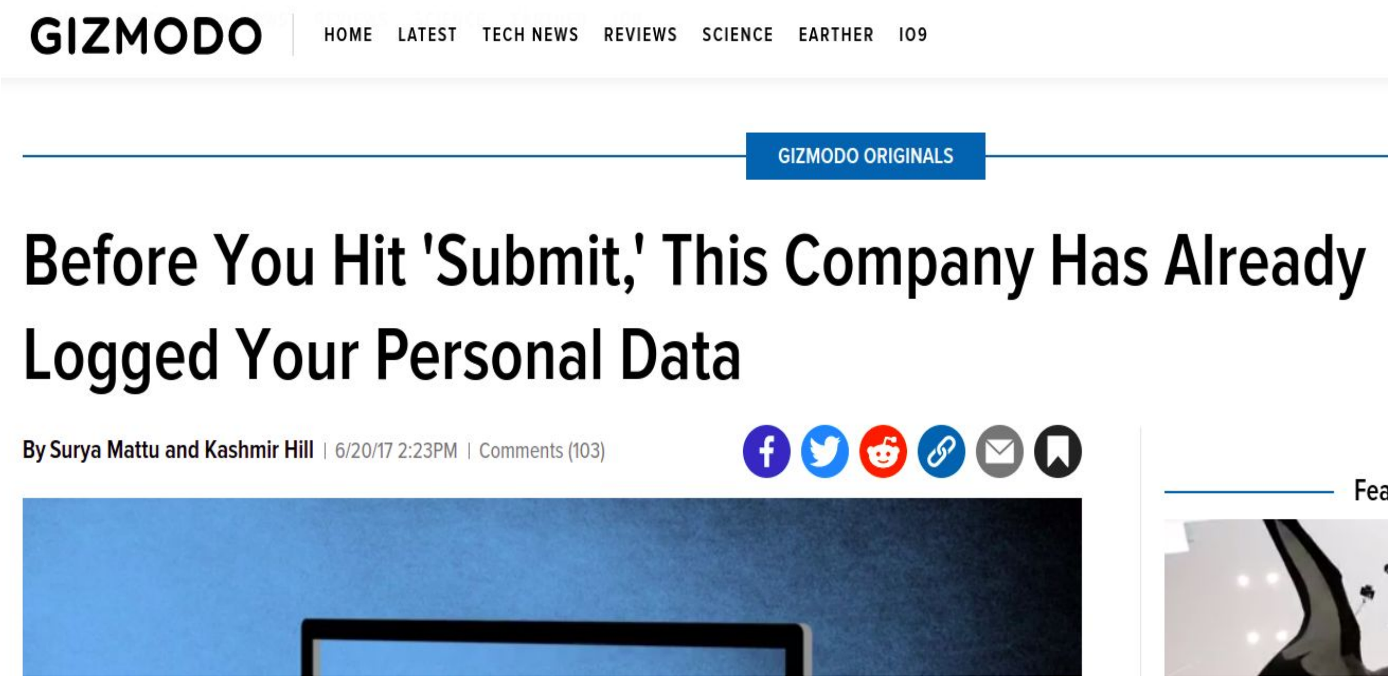
- Websites use advertising and marketing for monetization
  - built-in anti-tracking countermeasures
  - potential third-party cookie phase-out
- Tracking by email addresses
  - enables cross-site, cross-platform, persistent tracking



https://medium.com/@ugurekmekci/real-time-user-activity-tracking-w-divolte-collector-and-kafka-d8c106313400

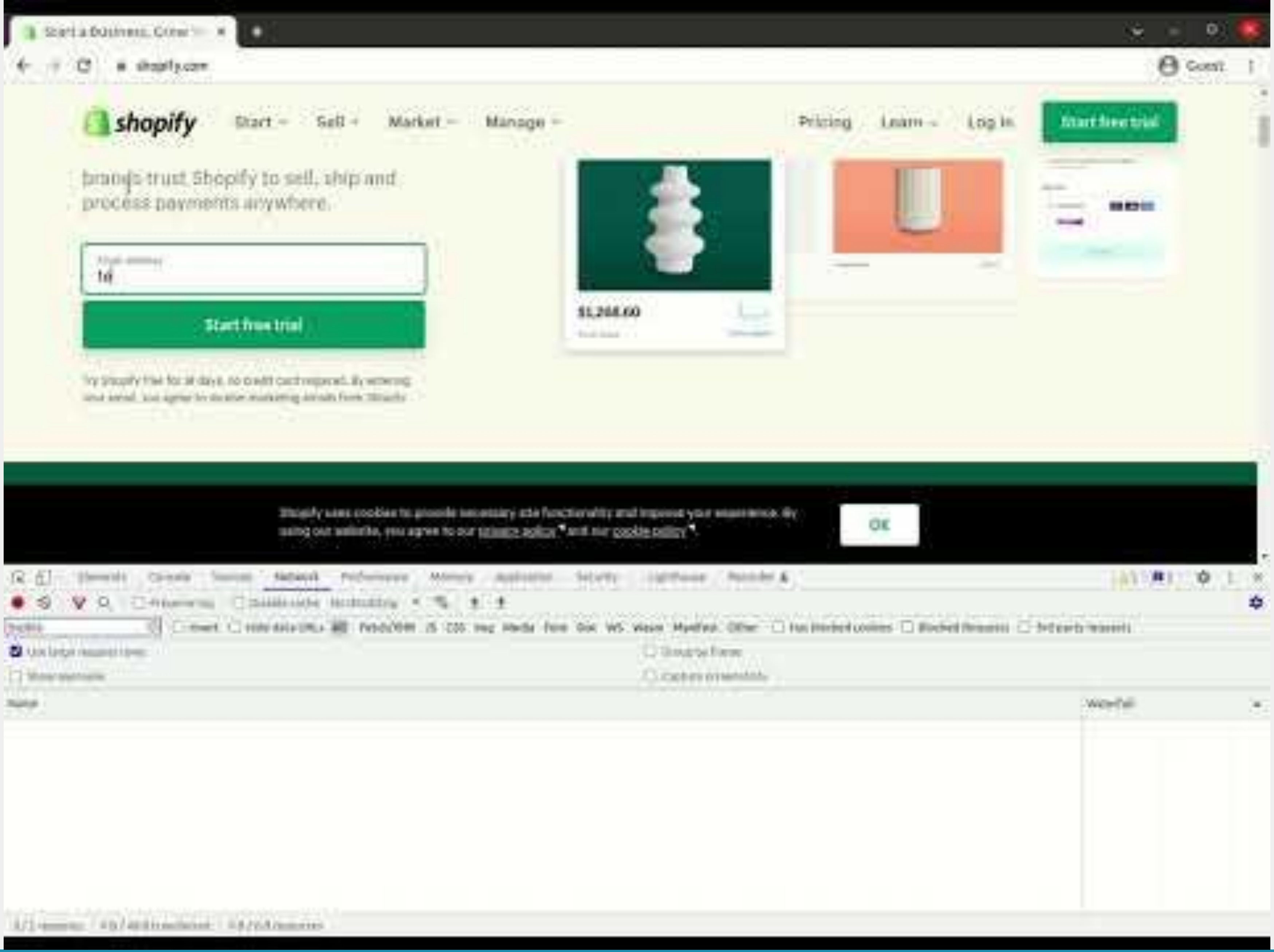
Motivation

- PII collection before form submission on a mortgage calculator website (Gizmodo, 2017)
- A 2018 survey (n=502):
  - 81% abandoned forms at least once
  - 59% abandoned a form in the last month [6].



Before You Hit 'Submit,' This Company Has Already Logged Your Personal Data

By Surya Mattu and Kashmir Hill | 6/20/17 2:23PM | Comments (103)



brands trust Shopify to sell, ship and process payments anywhere.

Email address
1q

Start free trial

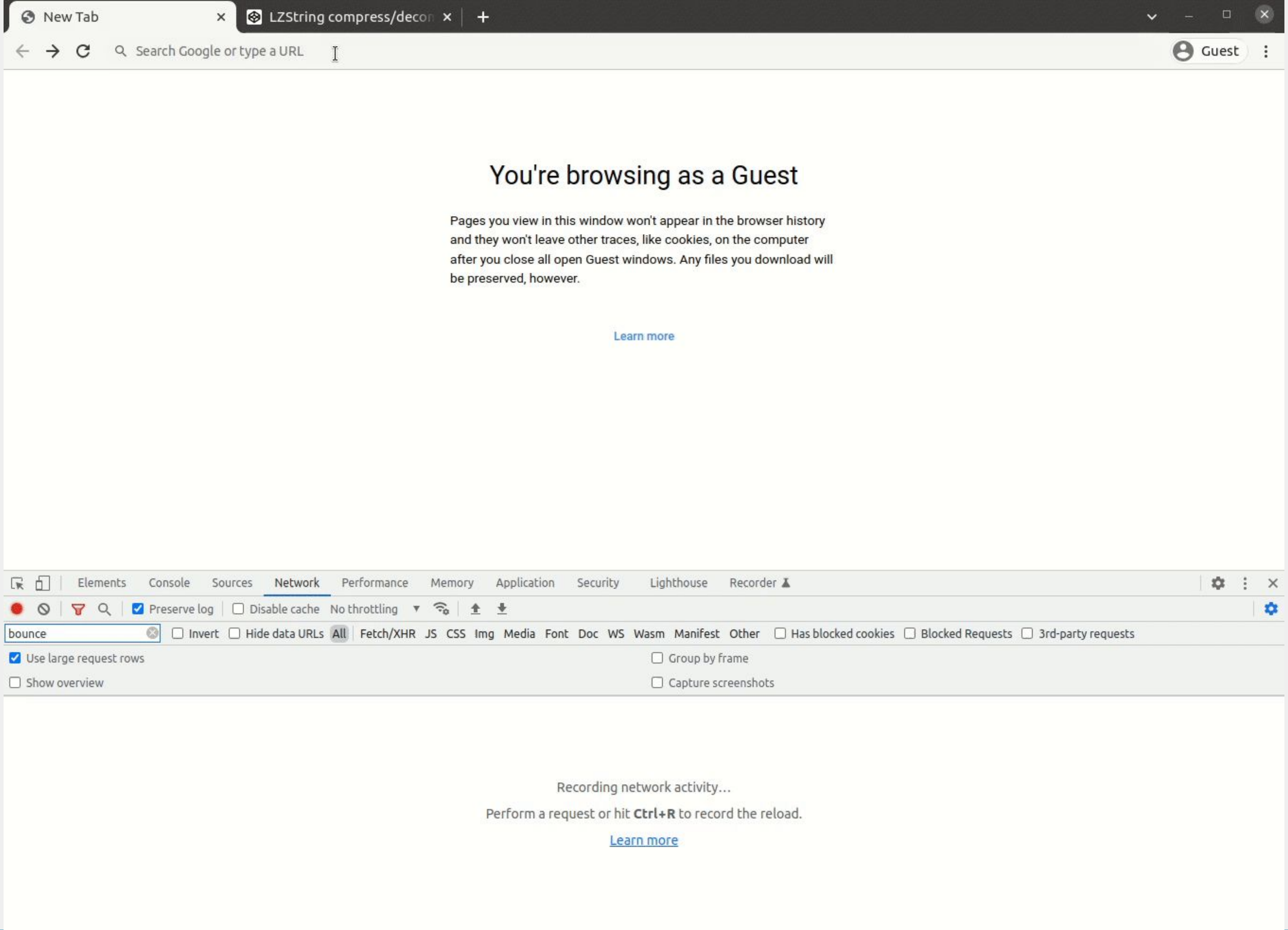
Try Shopify free for 14 days, no credit card required. By entering your email, you agree to receive marketing emails from Shopify.

Shopify uses cookies to provide necessary site functionality and improve your experience. By using our website, you agree to our privacy policy and our cookie policy.

OK

Objectives

- Measure email and password collection prior to form submission
    - effect of location: EU vs. US
    - effect of consent
    - mobile vs. desktop



You're browsing as a Guest

Pages you view in this window won't appear in the browser history and they won't leave other traces, like cookies, on the computer after you close all open Guest windows. Any files you download will be preserved, however.

Learn more

☑ Preserve log ☐ Disable cache No throttling

bounce ☐ Invert ☐ Hide data URLs All ☐ Has blocked cookies ☐ Blocked Requests ☐ 3rd-party requests

☑ Use large request rows ☐ Group by frame

☐ Show overview ☐ Capture screenshots

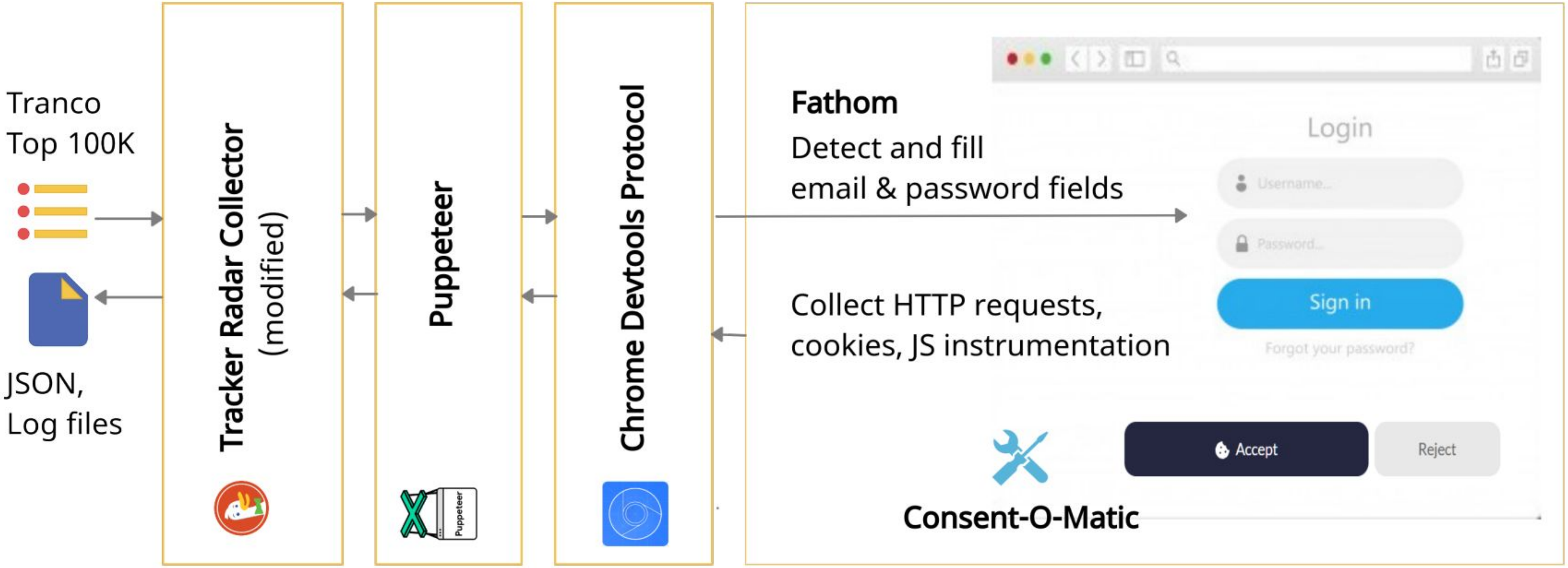
Recording network activity…

Perform a request or hit **Ctrl+R** to record the reload.

Learn more

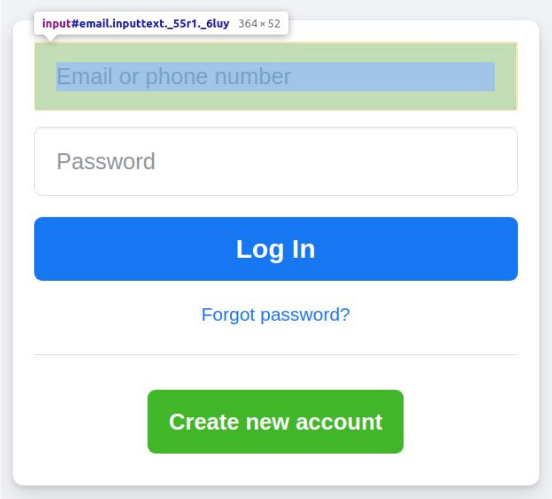
Method – Web Crawler

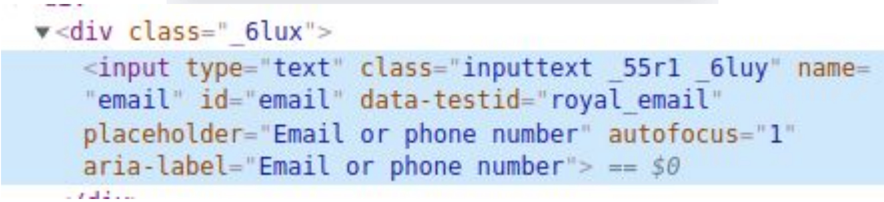
- Built on Tracker Radar Collector (developed by DuckDuckGo)



Method – Email field detection

- **Fathom**: A supervised learning framework specialized to detect webpage parts [8]

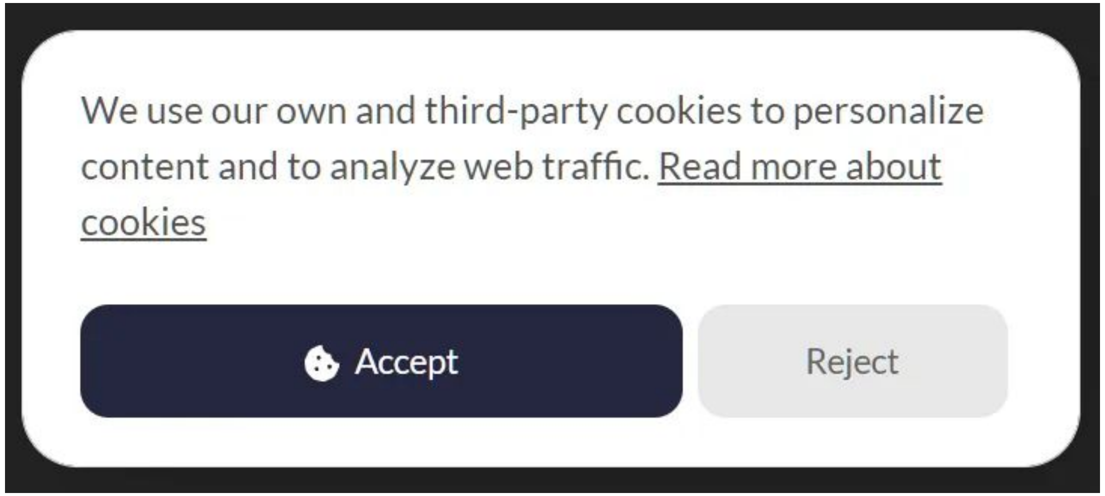




```
▼<div class="_6lux">
  <input type="text" class="inputtext _55r1 _6luy" name=
  "email" id="email" data-testid="royal_email"
  placeholder="Email or phone number" autofocus="1"
  aria-label="Email or phone number"> == $0
```

Method – CMP detection

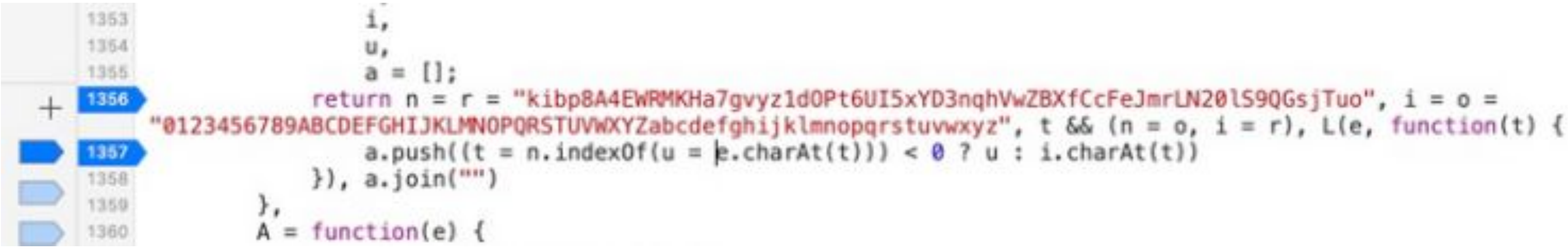
- **Consent-O-Matic**: browser extension that can recognize and interact with Consent Management Provider (CMP) pop-ups [7]
- Consent modes:
  - No-action
  - Accept-all
  - Reject-all
- ≈7,700/100K



We use our own and third-party cookies to personalize content and to analyze web traffic. Read more about cookies

Method – Leak Detection

- Based on Englehardt et al.'s method [3]
  - search for different encodings and hashes of search terms
- Identified two new encodings and a hashing method
  - LZString, custom mapping, hashing with a fixed salt



```
                i,
                u,
                a = [];
            return n = r = "kibp8A4EWRMKHa7gvyz1dOPt6UI5xYD3nqhVwZBXfCcFeJmrLN20lS9QGsjTuo", i = o =
"0123456789ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz", t && (n = o, i = r), L(e, function(t) {
                a.push((t = n.indexOf(u = e.charAt(t))) < 0 ? u : i.charAt(t))
            }), a.join("")
        },
        A = function(e) {
```

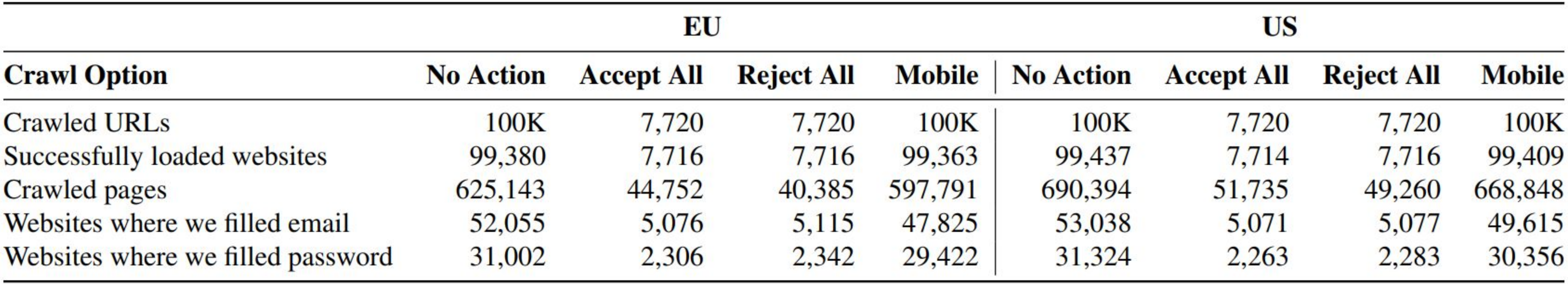
Method – Tracker Labeling

- Only considered leaks to tracker domains
- Block-lists we used:
  - Disconnect
  - Whotracks.me
  - DuckDuckGo (tds.json)
  - uBlock Origin
  - + Manual labeling

Crawls

- 8 Crawls, 100K websites, 2.8M pages, May & June 2021
- 2 locations: EU (Frankfurt), US (NYC)
- 3 consent modes:
  - no action (100K)
  - accept all (7,220)
  - reject all   (7,220)
- Mobile
  - EU, US
  - 100K – no action

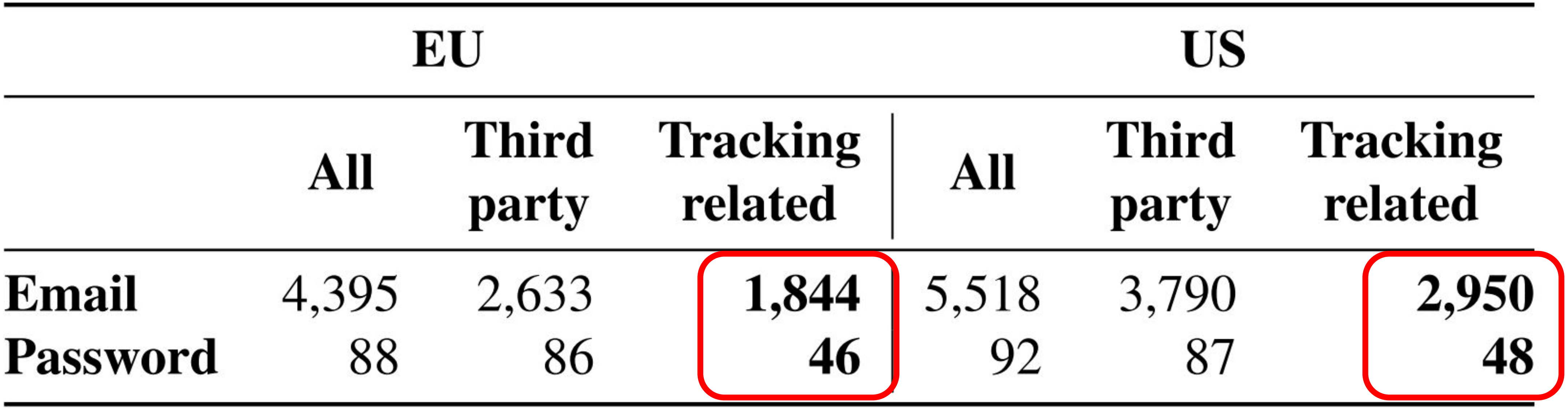
Dataset



	EU				US			
Crawl Option	**No Action**	**Accept All**	**Reject All**	**Mobile**	**No Action**	**Accept All**	**Reject All**	**Mobile**
Crawled URLs	100K	7,720	7,720	100K	100K	7,720	7,720	100K
Successfully loaded websites	99,380	7,716	7,716	99,363	99,437	7,714	7,716	99,409
Crawled pages	625,143	44,752	40,385	597,791	690,394	51,735	49,260	668,848
Websites where we filled email	52,055	5,076	5,115	47,825	53,038	5,071	5,077	49,615
Websites where we filled password	31,002	2,306	2,342	29,422	31,324	2,263	2,283	30,356

Overview of crawl statistics based on servers located in EU and the USA

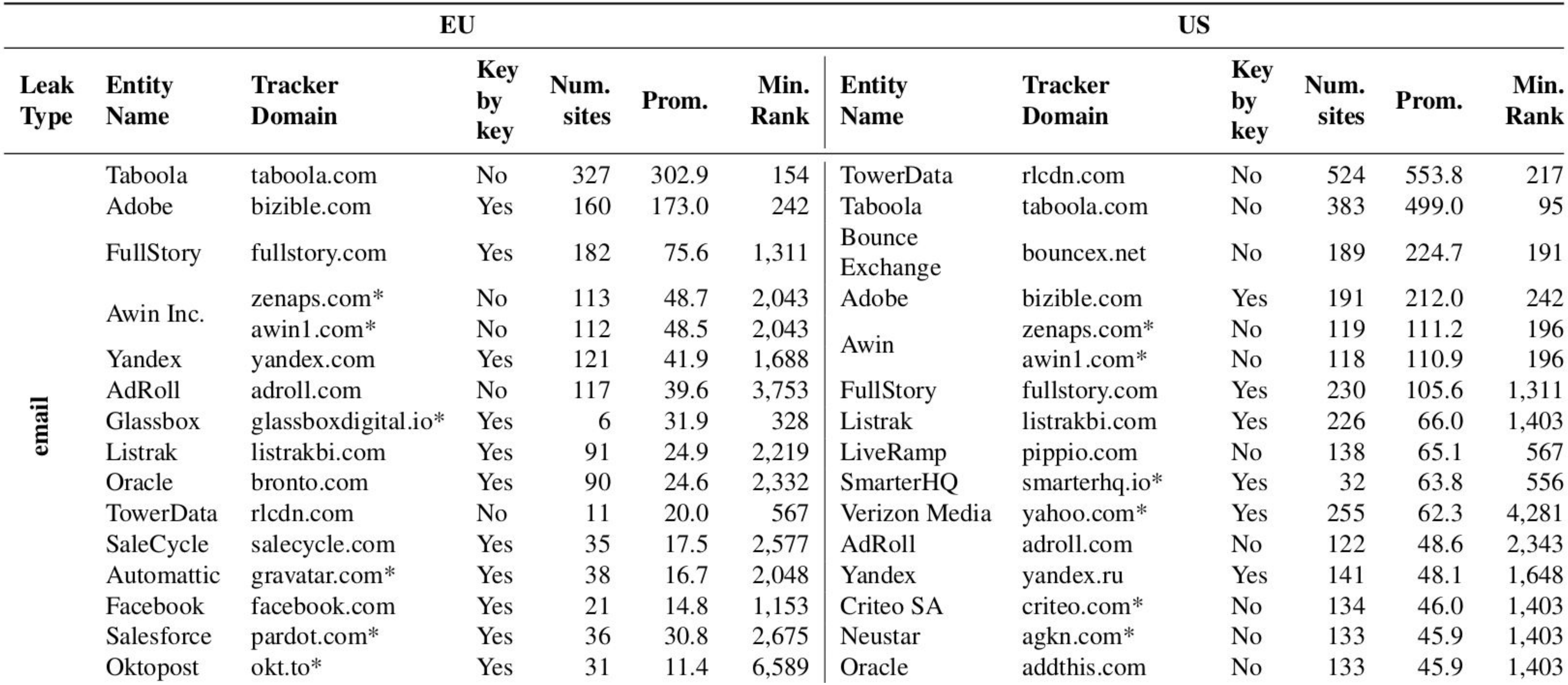
Results - Leaks



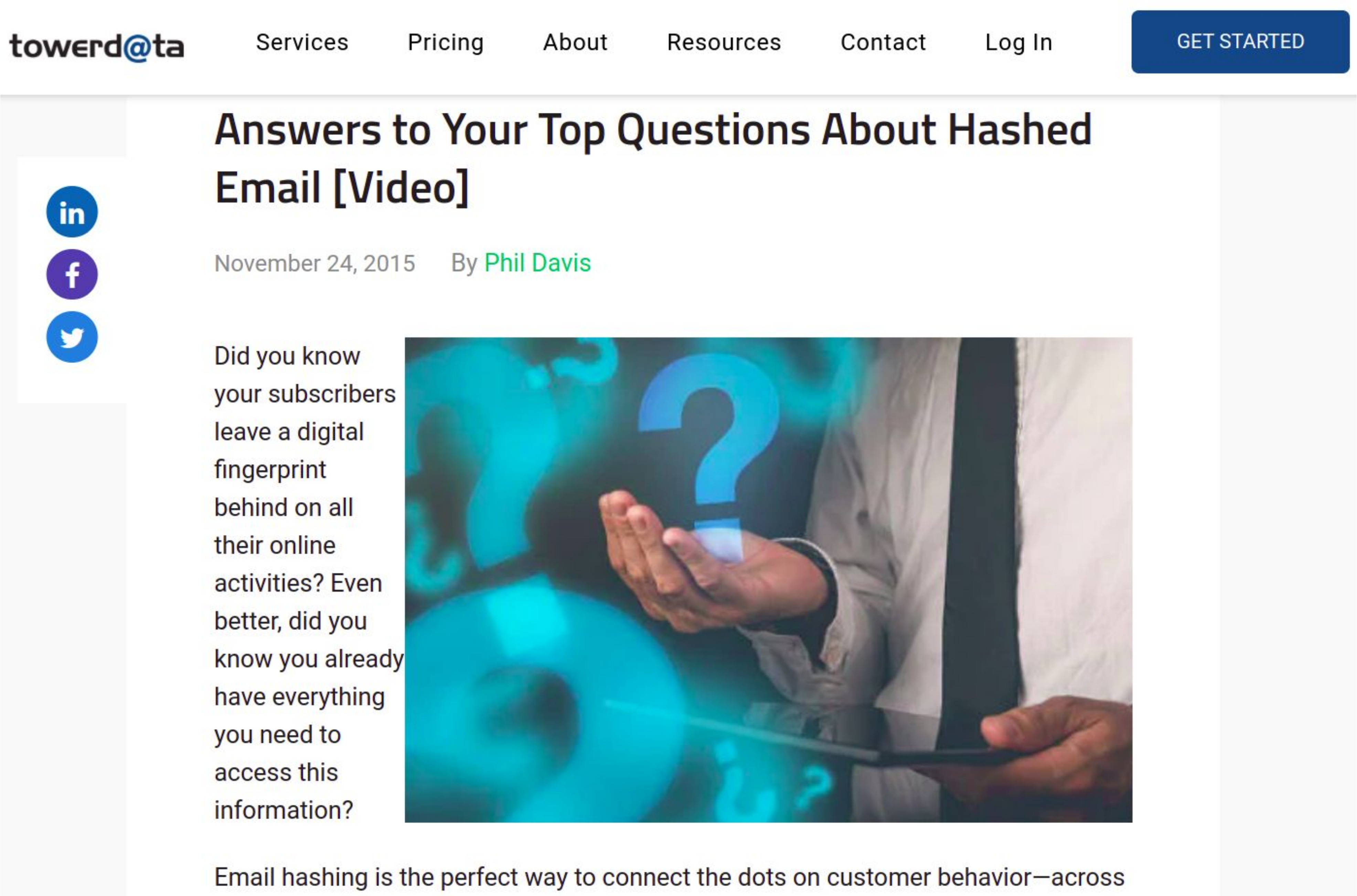
	EU			US		
	All	Third party	Tracking related	All	Third party	Tracking related
Email	4,395	2,633	**1,844**	5,518	3,790	**2,950**
Password	88	86	**46**	92	87	**48**

- Discovered 41 unlisted tracker domains

Results - Email Leaks



	EU						US					
Leak Type	Entity Name	Tracker Domain	Key by key	Num. sites	Prom.	Min. Rank	Entity Name	Tracker Domain	Key by key	Num. sites	Prom.	Min. Rank
email	Taboola	taboola.com	No	327	302.9	154	TowerData	rlcdn.com	No	524	553.8	217
	Adobe	bizible.com	Yes	160	173.0	242	Taboola	taboola.com	No	383	499.0	95
	FullStory	fullstory.com	Yes	182	75.6	1,311	Bounce Exchange	bouncex.net	No	189	224.7	191
	Awin Inc.	zenaps.com*	No	113	48.7	2,043	Adobe	bizible.com	Yes	191	212.0	242
		awin1.com*	No	112	48.5	2,043	Awin	zenaps.com*	No	119	111.2	196
	Yandex	yandex.com	Yes	121	41.9	1,688		awin1.com*	No	118	110.9	196
	AdRoll	adroll.com	No	117	39.6	3,753	FullStory	fullstory.com	Yes	230	105.6	1,311
	Glassbox	glassboxdigital.io*	Yes	6	31.9	328	Listrak	listrakbi.com	Yes	226	66.0	1,403
	Listrak	listrakbi.com	Yes	91	24.9	2,219	LiveRamp	pippio.com	No	138	65.1	567
	Oracle	bronto.com	Yes	90	24.6	2,332	SmarterHQ	smarterhq.io*	Yes	32	63.8	556
	TowerData	rlcdn.com	No	11	20.0	567	Verizon Media	yahoo.com*	Yes	255	62.3	4,281
	SaleCycle	salecycle.com	Yes	35	17.5	2,577	AdRoll	adroll.com	No	122	48.6	2,343
	Automattic	gravatar.com*	Yes	38	16.7	2,048	Yandex	yandex.ru	Yes	141	48.1	1,648
	Facebook	facebook.com	Yes	21	14.8	1,153	Criteo SA	criteo.com*	No	134	46.0	1,403
	Salesforce	pardot.com*	Yes	36	30.8	2,675	Neustar	agkn.com*	No	133	45.9	1,403
	Oktopost	okt.to*	Yes	31	11.4	6,589	Oracle	addthis.com	No	133	45.9	1,403

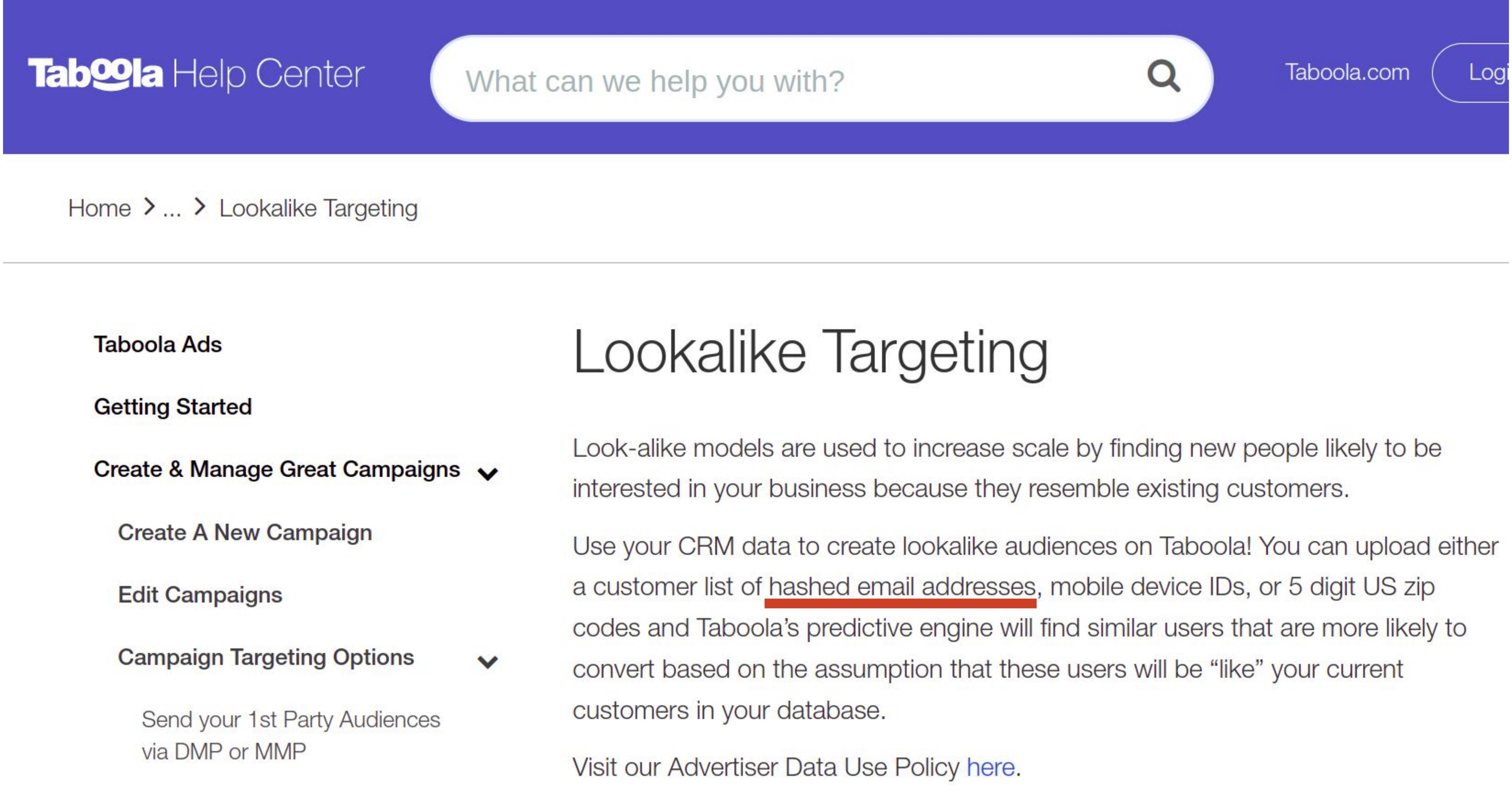


Answers to Your Top Questions About Hashed Email [Video]

November 24, 2015 By Phil Davis

Did you know your subscribers leave a digital fingerprint behind on all their online activities? Even better, did you know you already have everything you need to access this information?

Email hashing is the perfect way to connect the dots on customer behavior—across

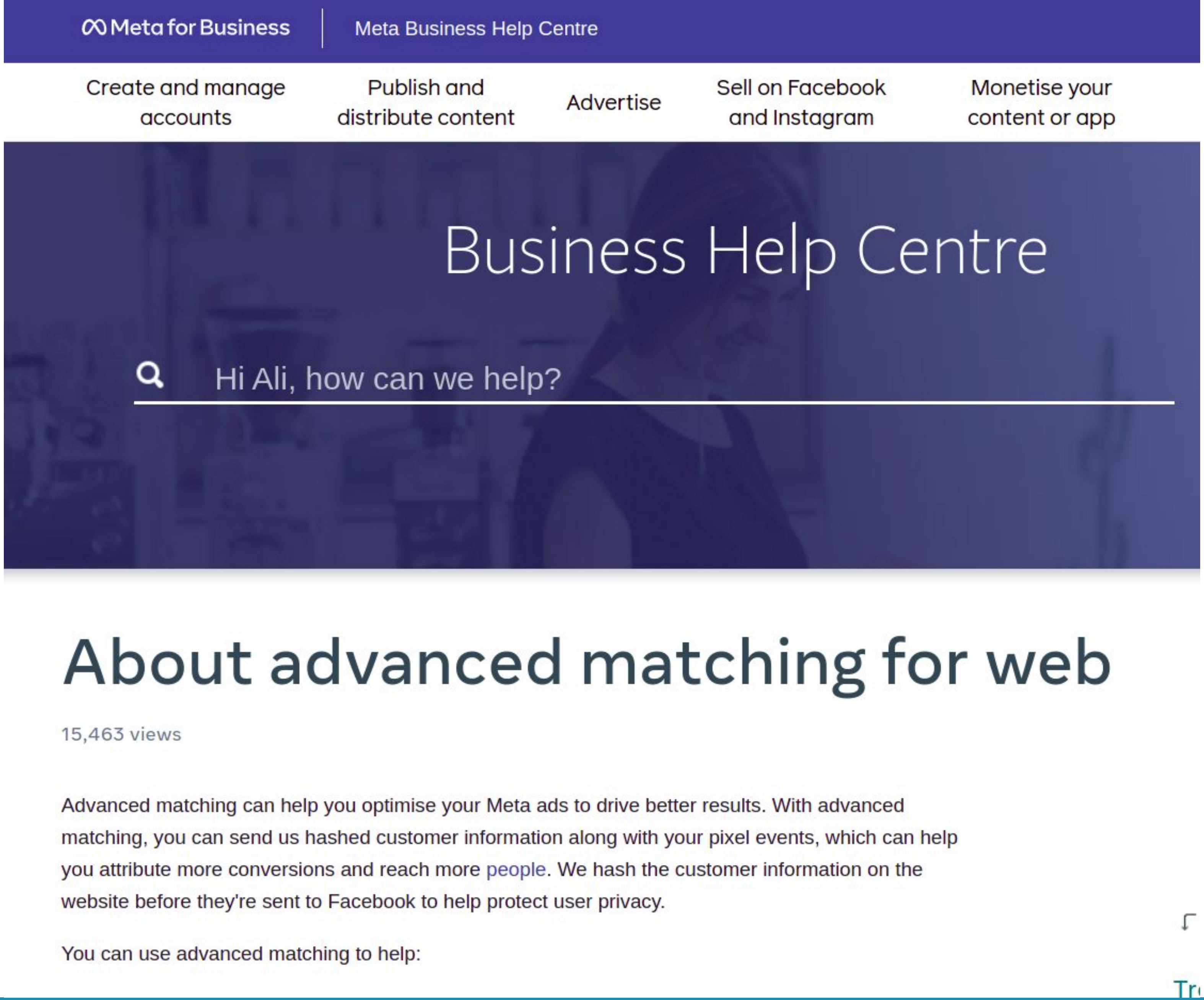

Taboola Help Center
What can we help you with?
Taboola.com
Logi
Home > ... > Lookalike Targeting
Taboola Ads
Getting Started
Create & Manage Great Campaigns
Create A New Campaign
Edit Campaigns
Campaign Targeting Options
Send your 1st Party Audiences via DMP or MMP

Lookalike Targeting

Look-alike models are used to increase scale by finding new people likely to be interested in your business because they resemble existing customers.

Use your CRM data to create lookalike audiences on Taboola! You can upload either a customer list of hashed email addresses, mobile device IDs, or 5 digit US zip codes and Taboola's predictive engine will find similar users that are more likely to convert based on the assumption that these users will be "like" your current customers in your database.

Visit our Advertiser Data Use Policy here.

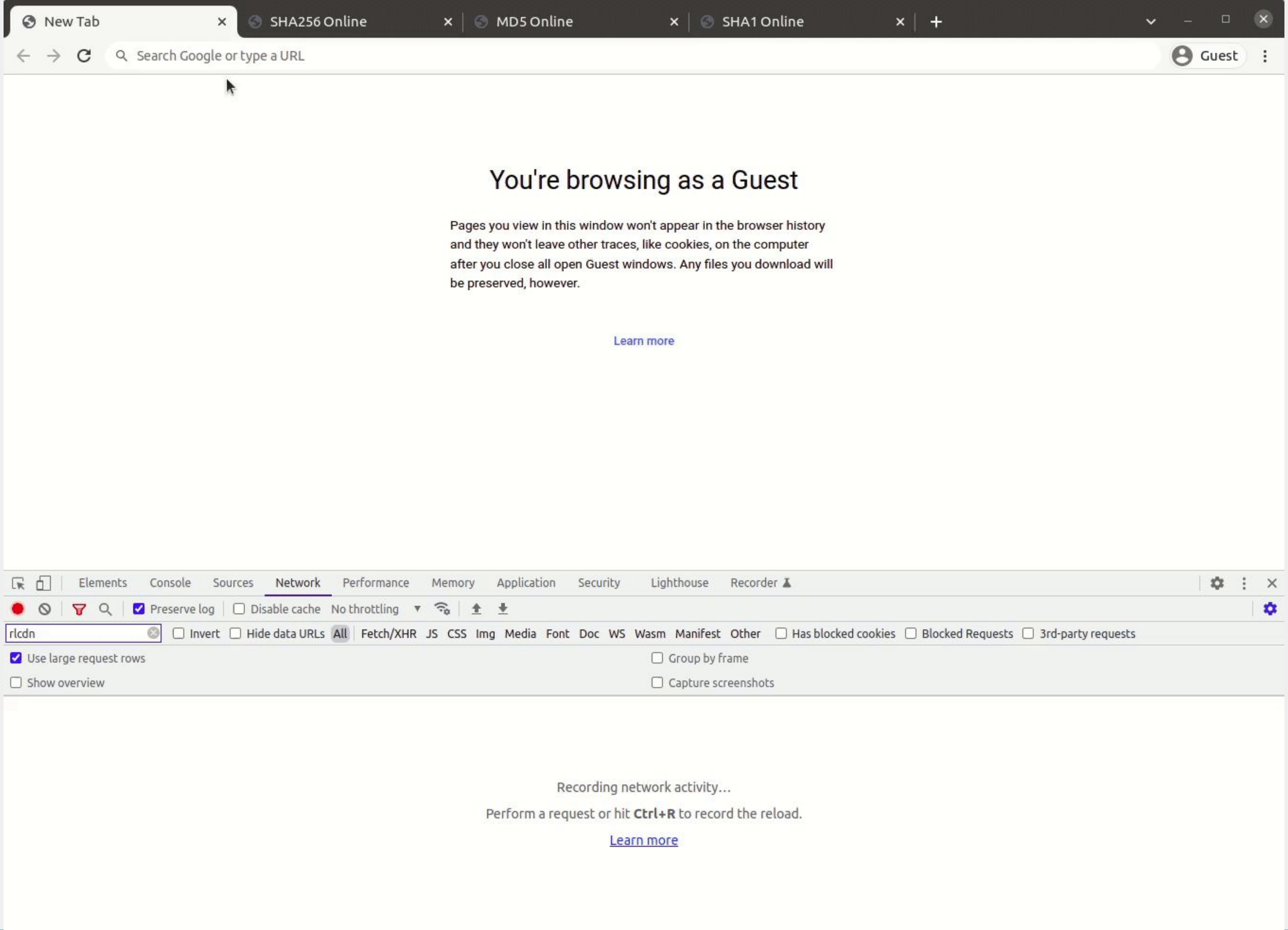

Meta for Business
Meta Business Help Centre
Create and manage accounts
Publish and distribute content
Advertise
Sell on Facebook and Instagram
Monetise your content or app
Business Help Centre
Hi Ali, how can we help?
Tr

About advanced matching for web

15,463 views

Advanced matching can help you optimise your Meta ads to drive better results. With advanced matching, you can send us hashed customer information along with your pixel events, which can help you attribute more conversions and reach more people. We hash the customer information on the website before they're sent to Facebook to help protect user privacy.

You can use advanced matching to help:



You're browsing as a Guest

Pages you view in this window won't appear in the browser history and they won't leave other traces, like cookies, on the computer after you close all open Guest windows. Any files you download will be preserved, however.

Learn more

☑ Preserve log ☐ Disable cache

rlcdn ☐ Invert ☐ Hide data URLs ☐ Has blocked cookies ☐ Blocked Requests ☐ 3rd-party requests

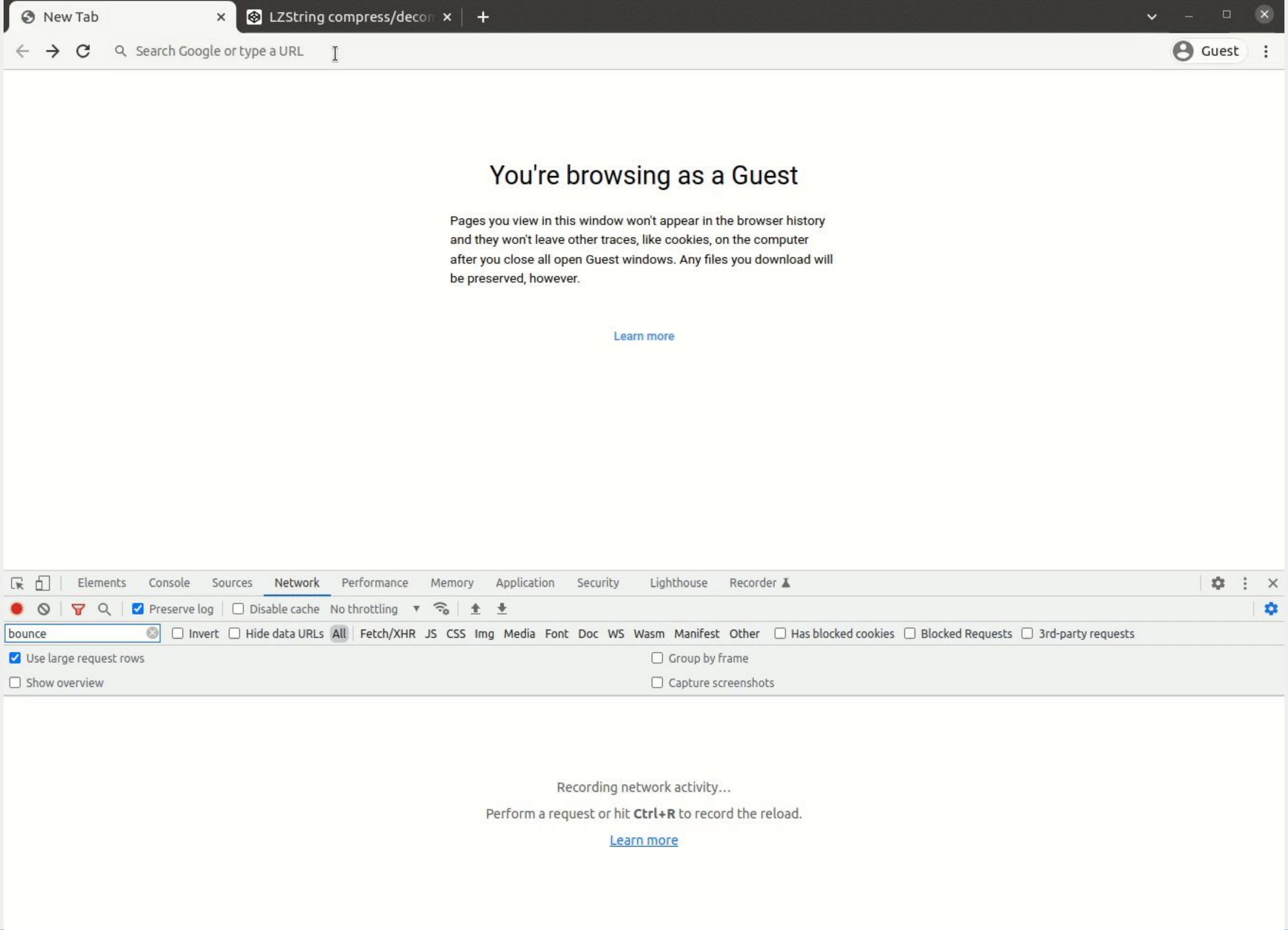
☑ Use large request rows ☐ Group by frame

☐ Show overview ☐ Capture screenshots

Recording network activity…

Perform a request or hit **Ctrl+R** to record the reload.

Learn more



You're browsing as a Guest

Pages you view in this window won't appear in the browser history and they won't leave other traces, like cookies, on the computer after you close all open Guest windows. Any files you download will be preserved, however.

Learn more

☑ Preserve log ☐ Disable cache

bounce ☐ Invert ☐ Hide data URLs ☐ Has blocked cookies ☐ Blocked Requests ☐ 3rd-party requests

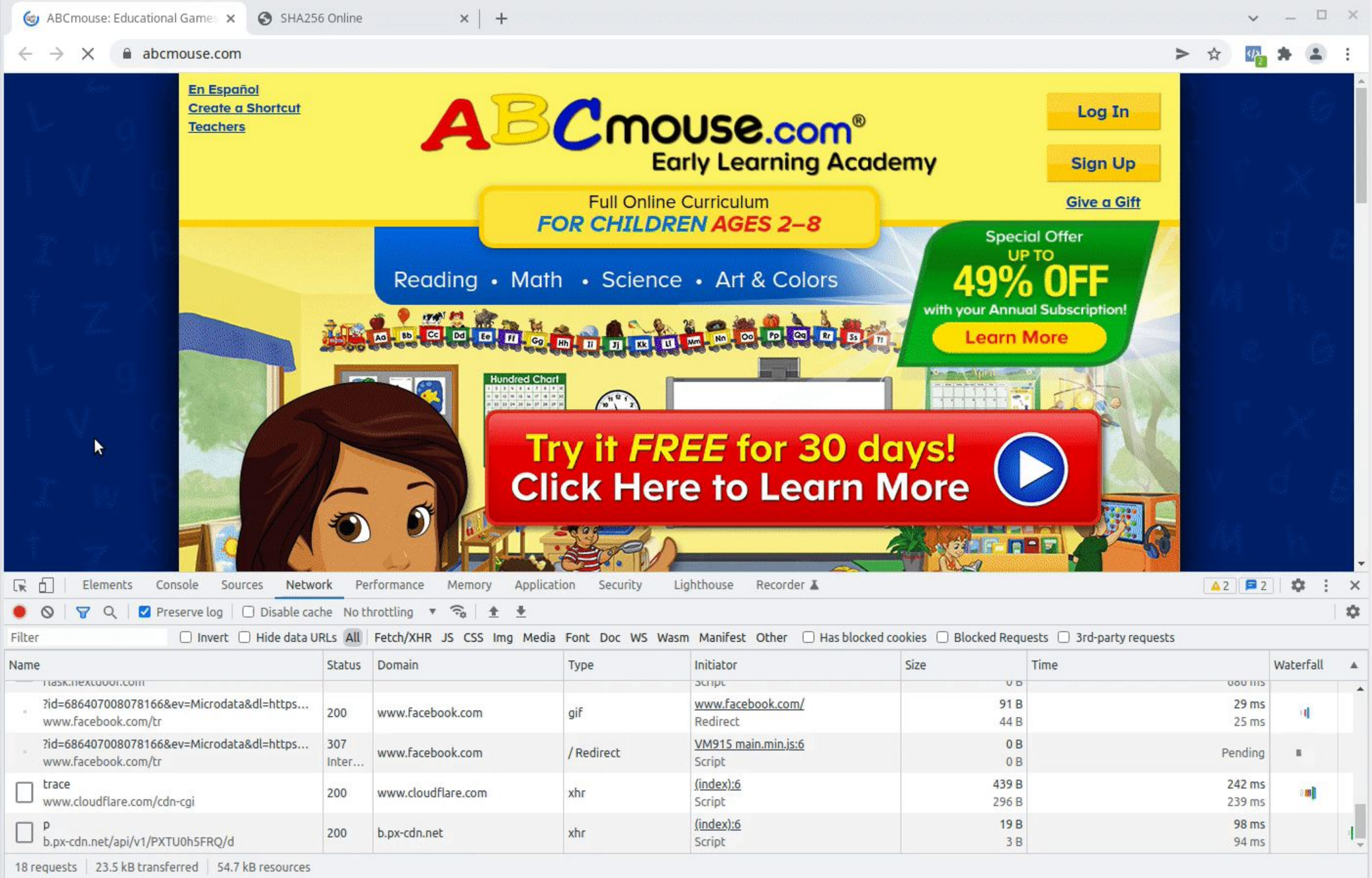
☑ Use large request rows ☐ Group by frame

☐ Show overview ☐ Capture screenshots

Recording network activity…

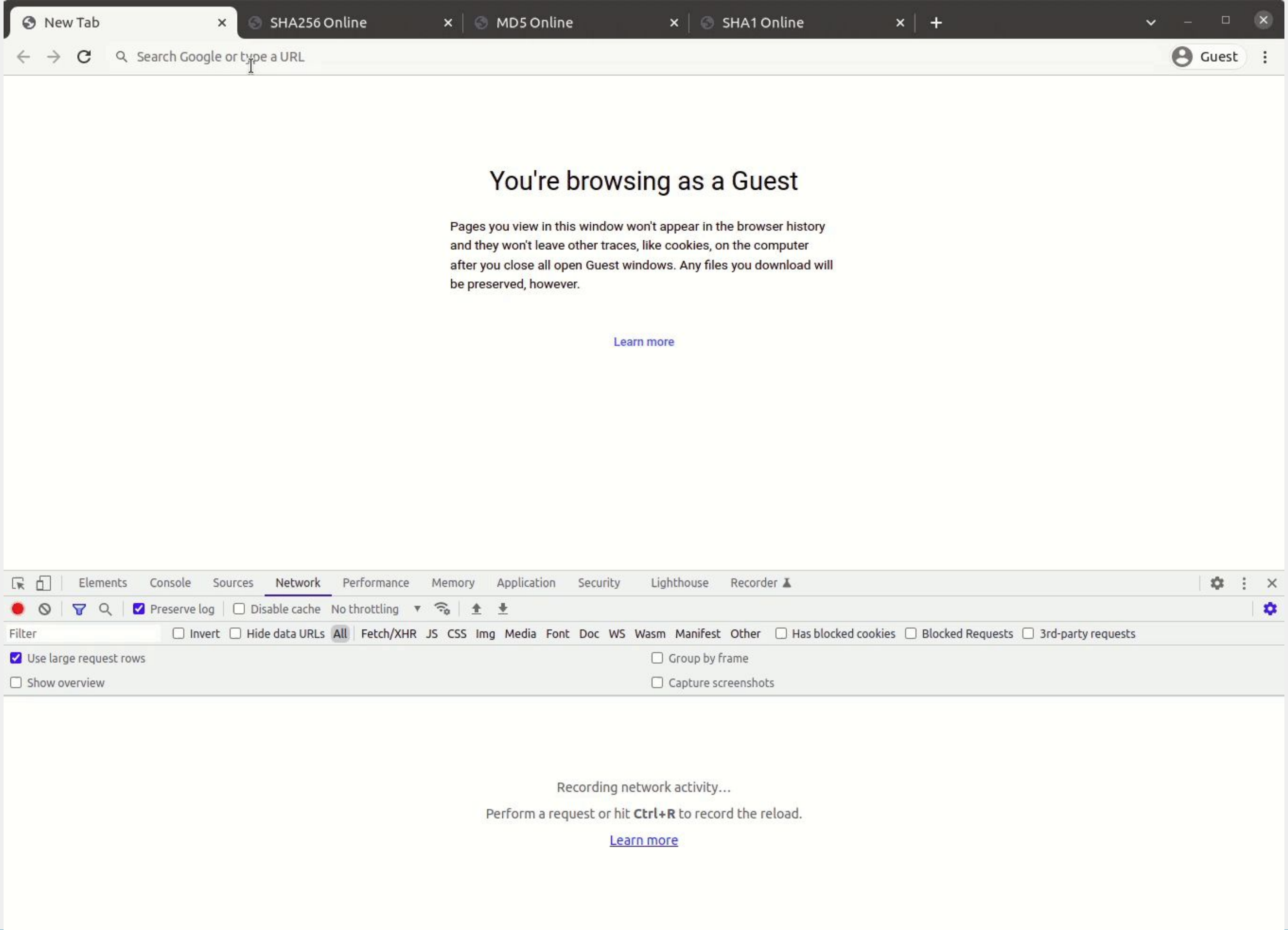
Perform a request or hit **Ctrl+R** to record the reload.

Learn more



Name	Status	Domain	Type	Initiator	Size	Time	Waterfall
?id=686407008078166&ev=Microdata&dl=https... www.facebook.com/tr	200	www.facebook.com	gif	www.facebook.com/ Redirect	91 B 44 B	29 ms 25 ms	
?id=686407008078166&ev=Microdata&dl=https... www.facebook.com/tr	307 Inter...	www.facebook.com	/ Redirect	VM915 main.min.js:6 Script	0 B 0 B	Pending	
trace www.cloudflare.com/cdn-cgi	200	www.cloudflare.com	xhr	(index):6 Script	439 B 296 B	242 ms 239 ms	
p b.px-cdn.net/api/v1/PXTU0h5FRQ/d	200	b.px-cdn.net	xhr	(index):6 Script	19 B 3 B	98 ms 94 ms	

18 requests | 23.5 kB transferred | 54.7 kB resources



You're browsing as a Guest

Pages you view in this window won't appear in the browser history and they won't leave other traces, like cookies, on the computer after you close all open Guest windows. Any files you download will be preserved, however.

Learn more

☑ Preserve log ☐ Disable cache No throttling

☐ Invert ☐ Hide data URLs ☐ Has blocked cookies ☐ Blocked Requests ☐ 3rd-party requests

☑ Use large request rows ☐ Group by frame

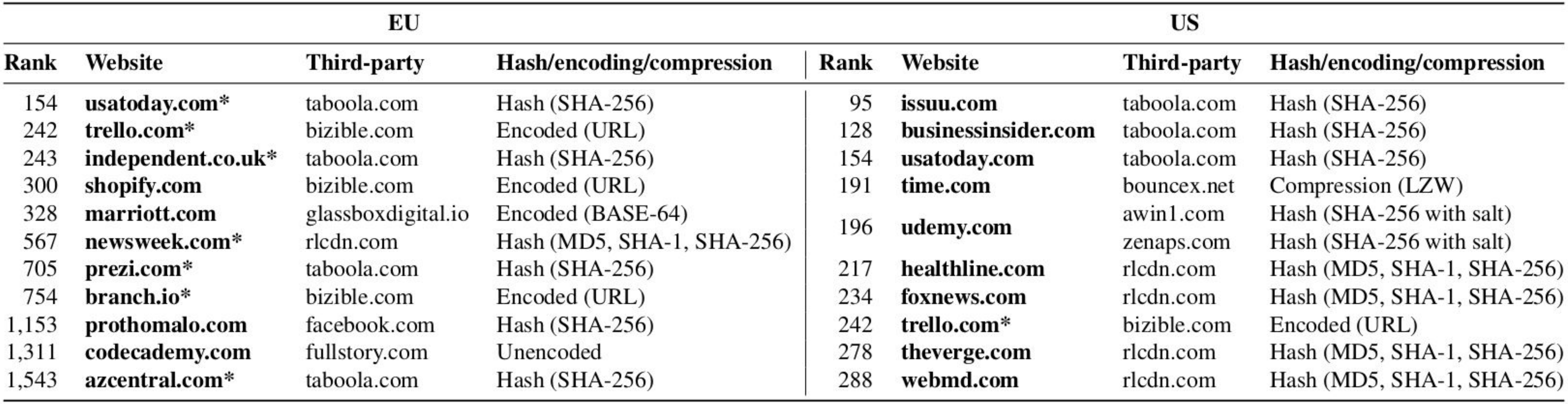
☐ Show overview ☐ Capture screenshots

Recording network activity…

Perform a request or hit **Ctrl+R** to record the reload.

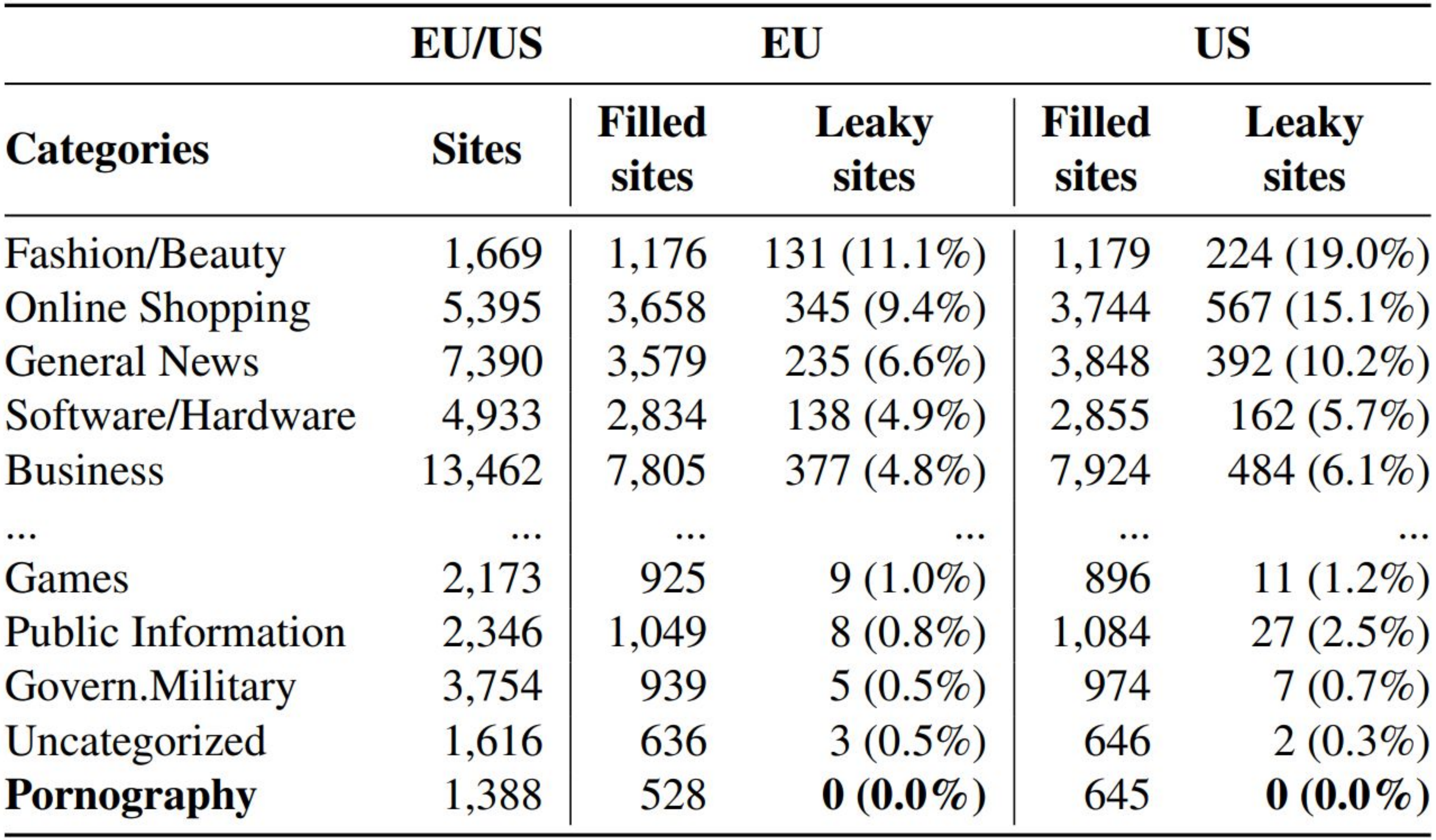
Learn more

Results - Top ten websites

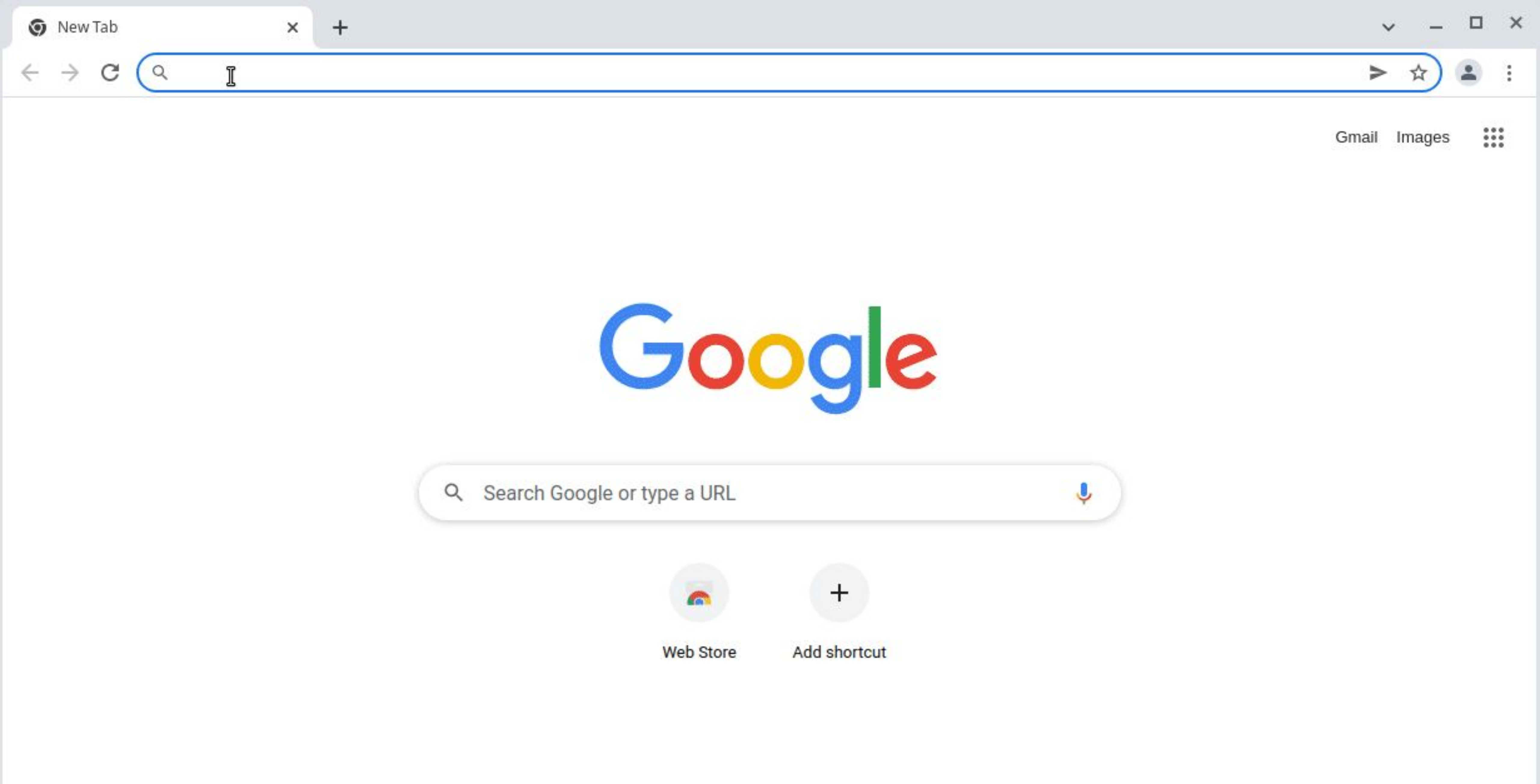


EU				US			
Rank	**Website**	**Third-party**	**Hash/encoding/compression**	**Rank**	**Website**	**Third-party**	**Hash/encoding/compression**
154	**usatoday.com***	taboola.com	Hash (SHA-256)	95	**issuu.com**	taboola.com	Hash (SHA-256)
242	**trello.com***	bizible.com	Encoded (URL)	128	**businessinsider.com**	taboola.com	Hash (SHA-256)
243	**independent.co.uk***	taboola.com	Hash (SHA-256)	154	**usatoday.com**	taboola.com	Hash (SHA-256)
300	**shopify.com**	bizible.com	Encoded (URL)	191	**time.com**	bouncex.net	Compression (LZW)
328	**marriott.com**	glassboxdigital.io	Encoded (BASE-64)	196	**udemy.com**	awin1.com	Hash (SHA-256 with salt)
567	**newsweek.com***	rlcdn.com	Hash (MD5, SHA-1, SHA-256)			zenaps.com	Hash (SHA-256 with salt)
705	**prezi.com***	taboola.com	Hash (SHA-256)	217	**healthline.com**	rlcdn.com	Hash (MD5, SHA-1, SHA-256)
754	**branch.io***	bizible.com	Encoded (URL)	234	**foxnews.com**	rlcdn.com	Hash (MD5, SHA-1, SHA-256)
1,153	**prothomalo.com**	facebook.com	Hash (SHA-256)	242	**trello.com***	bizible.com	Encoded (URL)
1,311	**codecademy.com**	fullstory.com	Unencoded	278	**theverge.com**	rlcdn.com	Hash (MD5, SHA-1, SHA-256)
1,543	**azcentral.com***	taboola.com	Hash (SHA-256)	288	**webmd.com**	rlcdn.com	Hash (MD5, SHA-1, SHA-256)

Results - Website Categories



	EU/US	EU		US	
Categories	**Sites**	**Filled sites**	**Leaky sites**	**Filled sites**	**Leaky sites**
Fashion/Beauty	1,669	1,176	131 (11.1%)	1,179	224 (19.0%)
Online Shopping	5,395	3,658	345 (9.4%)	3,744	567 (15.1%)
General News	7,390	3,579	235 (6.6%)	3,848	392 (10.2%)
Software/Hardware	4,933	2,834	138 (4.9%)	2,855	162 (5.7%)
Business	13,462	7,805	377 (4.8%)	7,924	484 (6.1%)
...	...	...	...	...	...
Games	2,173	925	9 (1.0%)	896	11 (1.2%)
Public Information	2,346	1,049	8 (0.8%)	1,084	27 (2.5%)
Govern.Military	3,754	939	5 (0.5%)	974	7 (0.7%)
Uncategorized	1,616	636	3 (0.5%)	646	2 (0.3%)
Pornography	1,388	528	**0 (0.0%)**	645	**0 (0.0%)**


New Tab
Gmail
Images
Google
Search Google or type a URL
Web Store
Add shortcut

Results - HTTP - WebSocket Usage

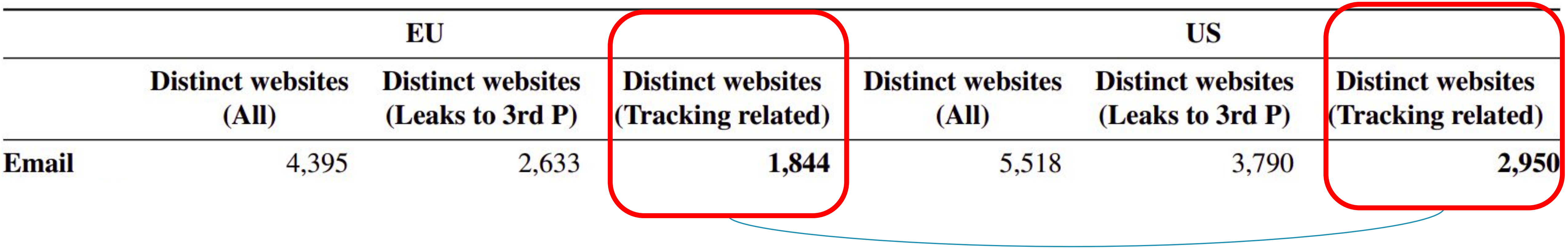
http://

- 15 websites in the EU
- 14 websites in the US

WebSocket

- To four tracker domains:
  - hotjar.com
  - freshrelevance.com
  - noibu.com
  - decibelinsight.net

Results - EU vs US



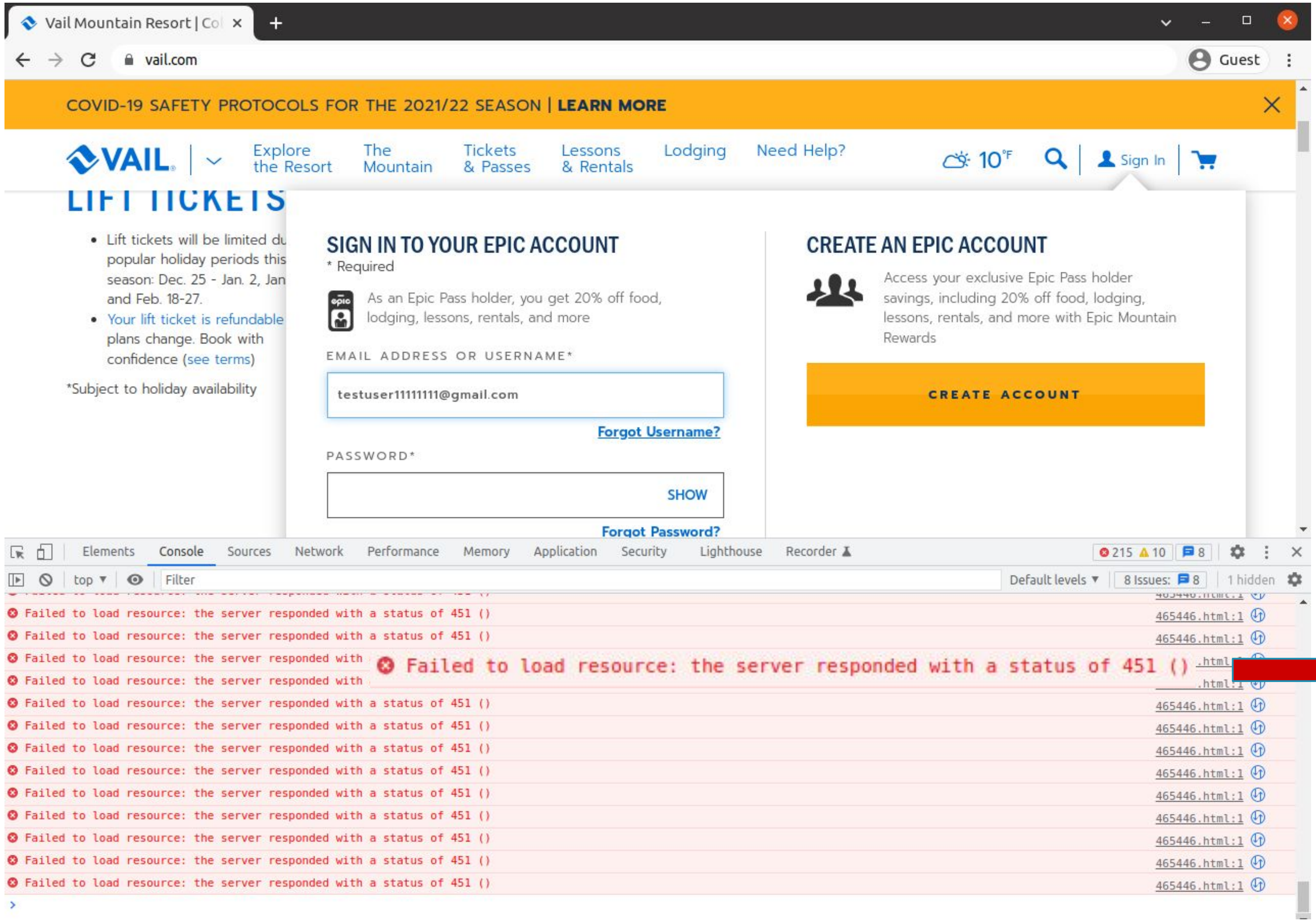
	EU			US		
	Distinct websites (All)	Distinct websites (Leaks to 3rd P)	Distinct websites (Tracking related)	Distinct websites (All)	Distinct websites (Leaks to 3rd P)	Distinct websites (Tracking related)
Email	4,395	2,633	**1,844**	5,518	3,790	**2,950**

60% difference

addthis.com, yahoo.com, doubleclick.net and criteo.com

Only appear in the US crawl

Results - EU vs US



COVID-19 SAFETY PROTOCOLS FOR THE 2021/22 SEASON | **LEARN MORE**

LIFT TICKETS

- Lift tickets will be limited du popular holiday periods this season: Dec. 25 - Jan. 2, Jan and Feb. 18-27.
- Your lift ticket is refundable plans change. Book with confidence (see terms)

*Subject to holiday availability

SIGN IN TO YOUR EPIC ACCOUNT

* Required

As an Epic Pass holder, you get 20% off food, lodging, lessons, rentals, and more

EMAIL ADDRESS OR USERNAME*

testuser11111111@gmail.com

Forgot Username?

PASSWORD*

SHOW

Forgot Password?

CREATE AN EPIC ACCOUNT

Access your exclusive Epic Pass holder savings, including 20% off food, lodging, lessons, rentals, and more with Epic Mountain Rewards

CREATE ACCOUNT

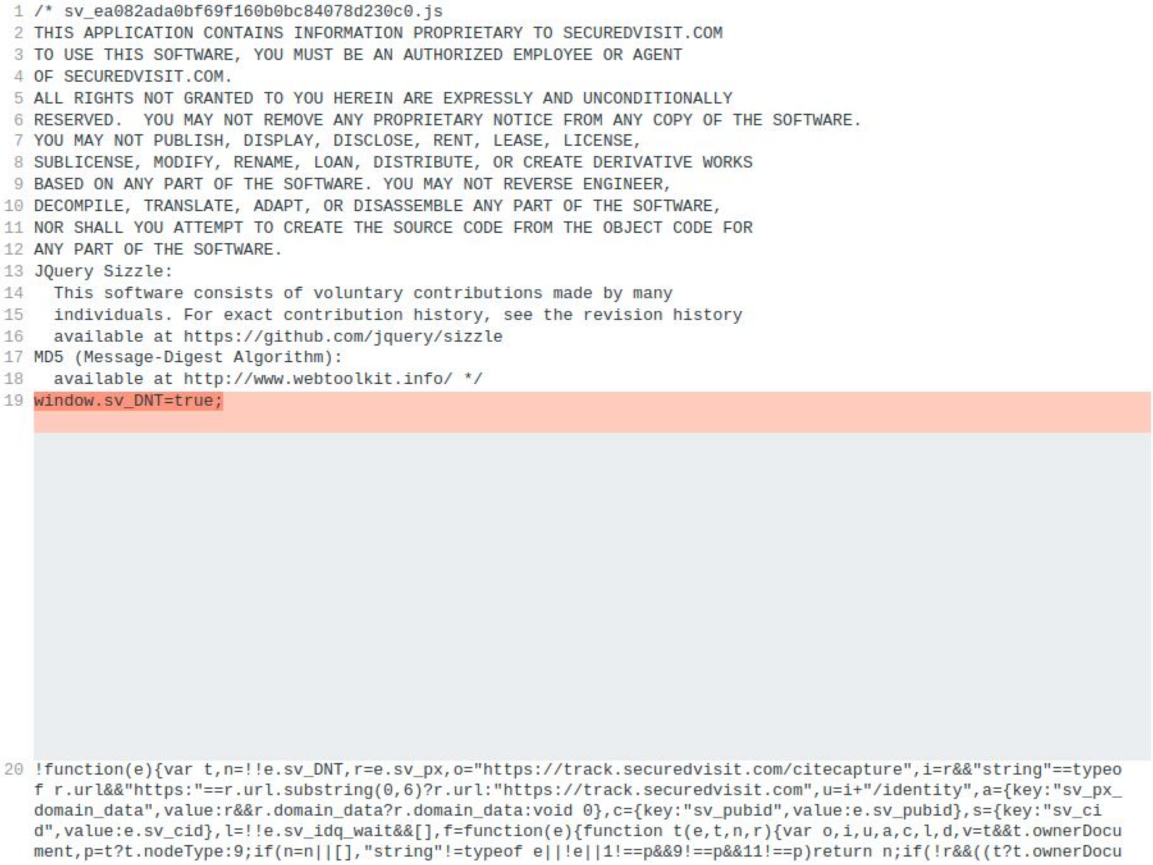
```
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with                       .html:1
Failed to load resource: the server responded with                       .html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
Failed to load resource: the server responded with a status of 451 ()    465446.html:1
```

```
Failed to load resource: the server responded with a status of 451 ()
```

rlcdn.com sends HTTP 451 error: Unavailable For Legal Reasons

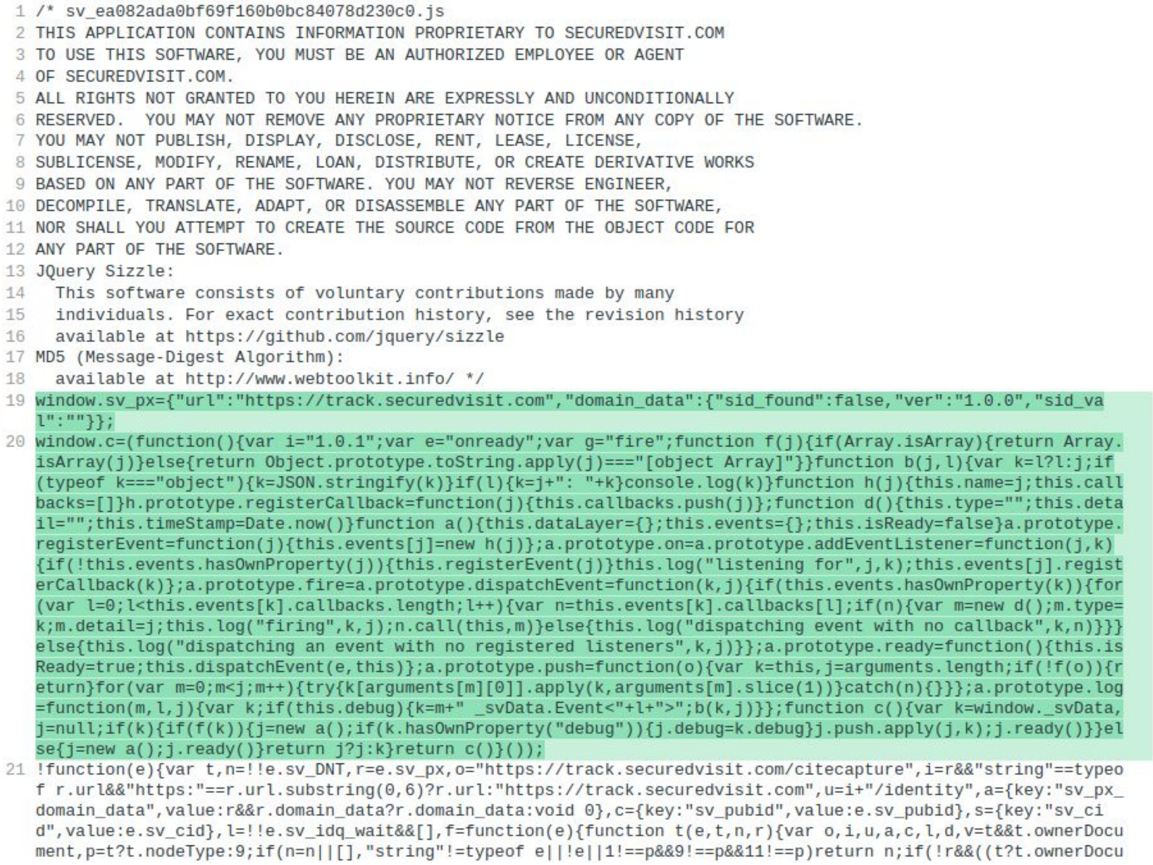
Results - EU vs US

Same script (from securedvisit.com) served with different content



```
/* sv_ea082ada0bf69f160b0bc84078d230c0.js
THIS APPLICATION CONTAINS INFORMATION PROPRIETARY TO SECUREDVISIT.COM
TO USE THIS SOFTWARE, YOU MUST BE AN AUTHORIZED EMPLOYEE OR AGENT
OF SECUREDVISIT.COM.
ALL RIGHTS NOT GRANTED TO YOU HEREIN ARE EXPRESSLY AND UNCONDITIONALLY
RESERVED.  YOU MAY NOT REMOVE ANY PROPRIETARY NOTICE FROM ANY COPY OF THE SOFTWARE.
YOU MAY NOT PUBLISH, DISPLAY, DISCLOSE, RENT, LEASE, LICENSE,
SUBLICENSE, MODIFY, RENAME, LOAN, DISTRIBUTE, OR CREATE DERIVATIVE WORKS
BASED ON ANY PART OF THE SOFTWARE. YOU MAY NOT REVERSE ENGINEER,
DECOMPILE, TRANSLATE, ADAPT, OR DISASSEMBLE ANY PART OF THE SOFTWARE,
NOR SHALL YOU ATTEMPT TO CREATE THE SOURCE CODE FROM THE OBJECT CODE FOR
ANY PART OF THE SOFTWARE.
JQuery Sizzle:
  This software consists of voluntary contributions made by many
  individuals. For exact contribution history, see the revision history
  available at https://github.com/jquery/sizzle
MD5 (Message-Digest Algorithm):
  available at http://www.webtoolkit.info/ */
window.sv_DNT=true;
!function(e){var t,n=!!e.sv_DNT,r=e.sv_px,o="https://track.securedvisit.com/citecapture",i=r&&"string"==typeo
f r.url&&"https:"==r.url.substring(0,6)?r.url:"https://track.securedvisit.com",u=i+"/identity",a={key:"sv_px_
domain_data",value:r&&r.domain_data?r.domain_data:void 0},c={key:"sv_pubid",value:e.sv_pubid},s={key:"sv_ci
d",value:e.sv_cid},l=!!e.sv_idq_wait&&[],f=function(e){function t(e,t,n,r){var o,i,u,a,c,l,d,v=t&&t.ownerDocu
ment,p=t?t.nodeType:9;if(n=n||[],"string"!=typeof e||!e||1!==p&&9!==p&&11!==p)return n;if(!r&&((t?t.ownerDocu
```

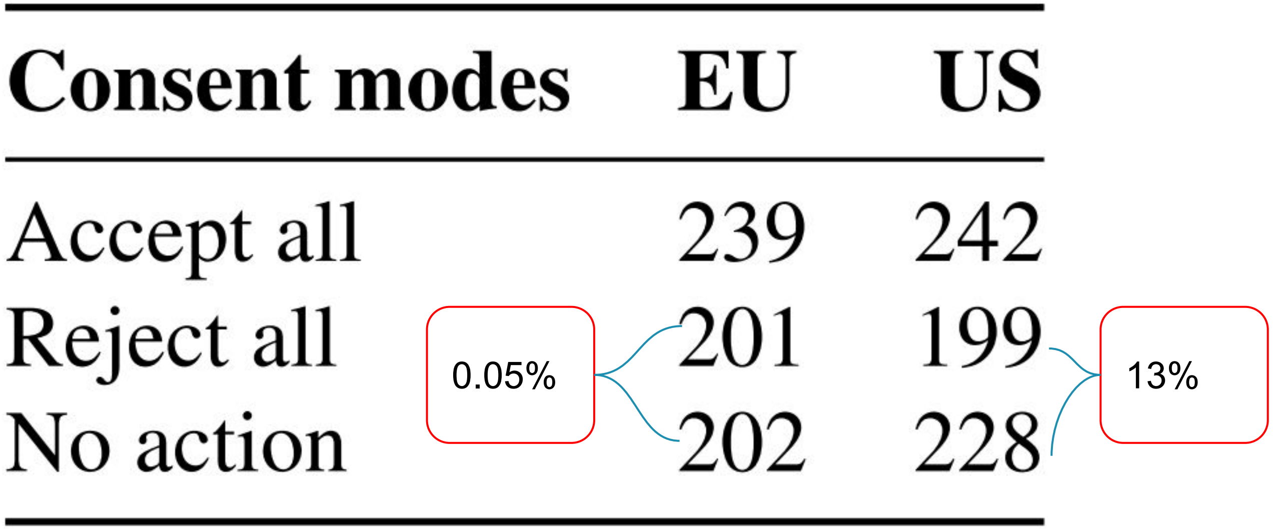
in the EU



```
/* sv_ea082ada0bf69f160b0bc84078d230c0.js
THIS APPLICATION CONTAINS INFORMATION PROPRIETARY TO SECUREDVISIT.COM
TO USE THIS SOFTWARE, YOU MUST BE AN AUTHORIZED EMPLOYEE OR AGENT
OF SECUREDVISIT.COM.
ALL RIGHTS NOT GRANTED TO YOU HEREIN ARE EXPRESSLY AND UNCONDITIONALLY
RESERVED.  YOU MAY NOT REMOVE ANY PROPRIETARY NOTICE FROM ANY COPY OF THE SOFTWARE.
YOU MAY NOT PUBLISH, DISPLAY, DISCLOSE, RENT, LEASE, LICENSE,
SUBLICENSE, MODIFY, RENAME, LOAN, DISTRIBUTE, OR CREATE DERIVATIVE WORKS
BASED ON ANY PART OF THE SOFTWARE. YOU MAY NOT REVERSE ENGINEER,
DECOMPILE, TRANSLATE, ADAPT, OR DISASSEMBLE ANY PART OF THE SOFTWARE,
NOR SHALL YOU ATTEMPT TO CREATE THE SOURCE CODE FROM THE OBJECT CODE FOR
ANY PART OF THE SOFTWARE.
JQuery Sizzle:
  This software consists of voluntary contributions made by many
  individuals. For exact contribution history, see the revision history
  available at https://github.com/jquery/sizzle
MD5 (Message-Digest Algorithm):
  available at http://www.webtoolkit.info/ */
window.sv_px={"url":"https://track.securedvisit.com","domain_data":{"sid_found":false,"ver":"1.0.0","sid_va
l":""}};
window.c=(function(){var i="1.0.1";var e="onready";var g="fire";function f(j){if(Array.isArray){return Array.
isArray(j)}else{return Object.prototype.toString.apply(j)==="[object Array]"}}function b(j,l){var k=l?l:j;if
(typeof k==="object"){k=JSON.stringify(k)}if(l){k=j+": "+k}console.log(k)}function h(j){this.name=j;this.call
backs=[]}h.prototype.registerCallback=function(j){this.callbacks.push(j)};function d(){this.type="";this.deta
il="";this.timeStamp=Date.now()}function a(){this.dataLayer={};this.events={};this.isReady=false}a.prototype.
registerEvent=function(j){this.events[j]=new h(j)};a.prototype.on=a.prototype.addEventListener=function(j,k)
{if(!this.events.hasOwnProperty(j)){this.registerEvent(j)}this.log("listening for",j,k);this.events[j].regist
erCallback(k)};a.prototype.fire=a.prototype.dispatchEvent=function(k,j){if(this.events.hasOwnProperty(k)){for
(var l=0;l<this.events[k].callbacks.length;l++){var n=this.events[k].callbacks[l];if(n){var m=new d();m.type=
k;m.detail=j;this.log("firing",k,j);n.call(this,m)}else{this.log("dispatching event with no callback",k,n)}}}
else{this.log("dispatching an event with no registered listeners",k,j)}};a.prototype.ready=function(){this.is
Ready=true;this.dispatchEvent(e,this)};a.prototype.push=function(o){var k=this,j=arguments.length;if(!f(o)){r
eturn}for(var m=0;m<j;m++){try{k[arguments[m][0]].apply(k,arguments[m].slice(1))}catch(n){}}};a.prototype.log
=function(m,l,j){var k;if(this.debug){k=m+" _svData.Event<"+l+">";b(k,j)}};function c(){var k=window._svData,
j=null;if(k){if(f(k)){j=new a();if(k.hasOwnProperty("debug")){j.debug=k.debug}j.push.apply(j,k);j.ready()}}el
se{j=new a();j.ready()}return j?j:k}return c()}());
!function(e){var t,n=!!e.sv_DNT,r=e.sv_px,o="https://track.securedvisit.com/citecapture",i=r&&"string"==typeo
f r.url&&"https:"==r.url.substring(0,6)?r.url:"https://track.securedvisit.com",u=i+"/identity",a={key:"sv_px_
domain_data",value:r&&r.domain_data?r.domain_data:void 0},c={key:"sv_pubid",value:e.sv_pubid},s={key:"sv_ci
d",value:e.sv_cid},l=!!e.sv_idq_wait&&[],f=function(e){function t(e,t,n,r){var o,i,u,a,c,l,d,v=t&&t.ownerDocu
ment,p=t?t.nodeType:9;if(n=n||[],"string"!=typeof e||!e||1!==p&&9!==p&&11!==p)return n;if(!r&&((t?t.ownerDocu
```

in the US

Results - The Effect of Consent

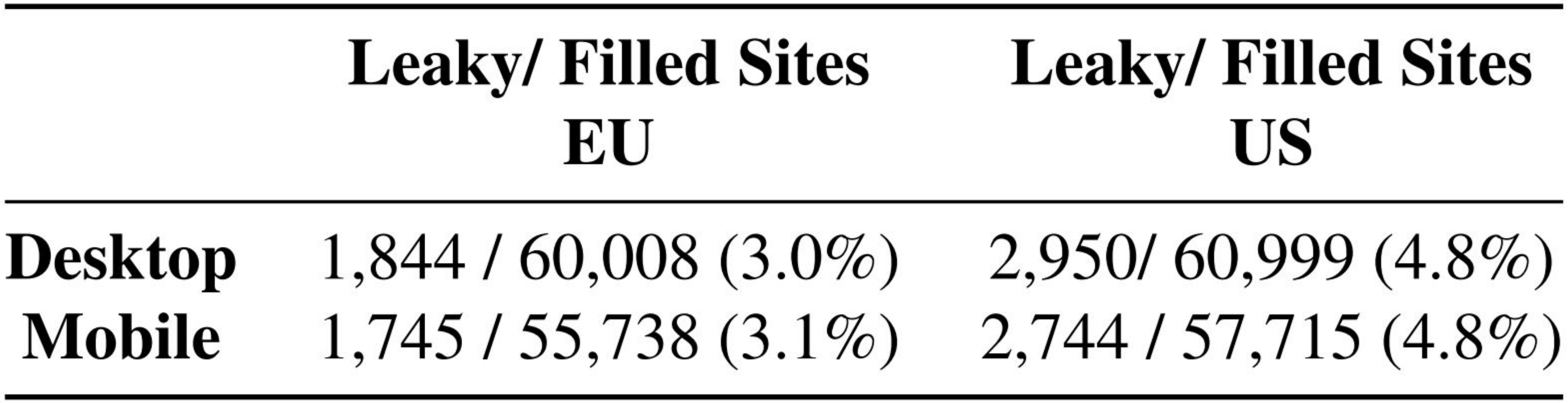


Consent modes	EU	US
Accept all	239	242
Reject all	201	199
No action	202	228

0.05%

13%

Results - Mobile

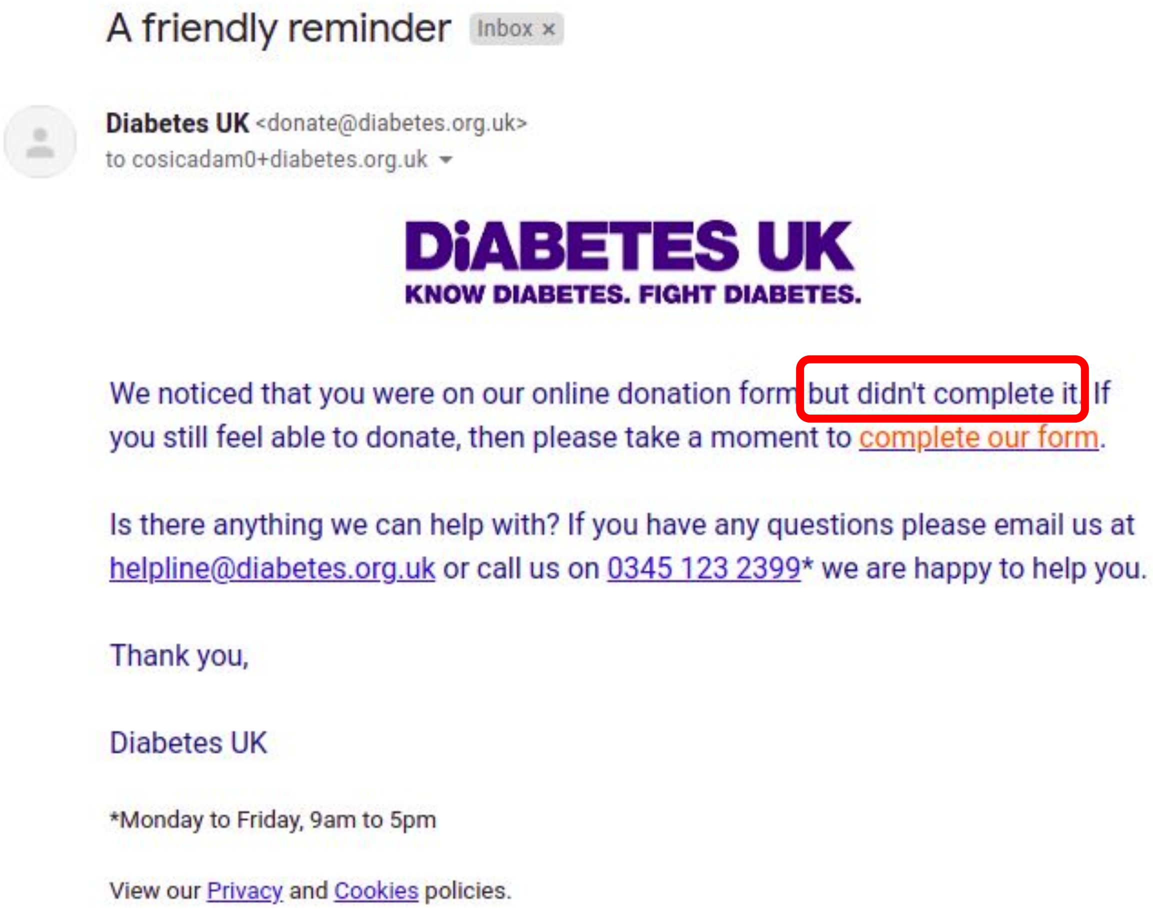


	Leaky/ Filled Sites EU	Leaky/ Filled Sites US
Desktop	1,844 / 60,008 (3.0%)	2,950/ 60,999 (4.8%)
Mobile	1,745 / 55,738 (3.1%)	2,744 / 57,715 (4.8%)

Results - Received Emails

- 290 emails from 88 distinct sites

Email from: diabetes.org.uk
Tracker domain: freshaddress.biz



A friendly reminder Inbox ×

Diabetes UK <donate@diabetes.org.uk>
to cosicadam0+diabetes.org.uk

DiABETES UK
KNOW DIABETES. FIGHT DIABETES.

We noticed that you were on our online donation form but didn't complete it. If you still feel able to donate, then please take a moment to complete our form.

Is there anything we can help with? If you have any questions please email us at helpline@diabetes.org.uk or call us on 0345 123 2399* we are happy to help you.

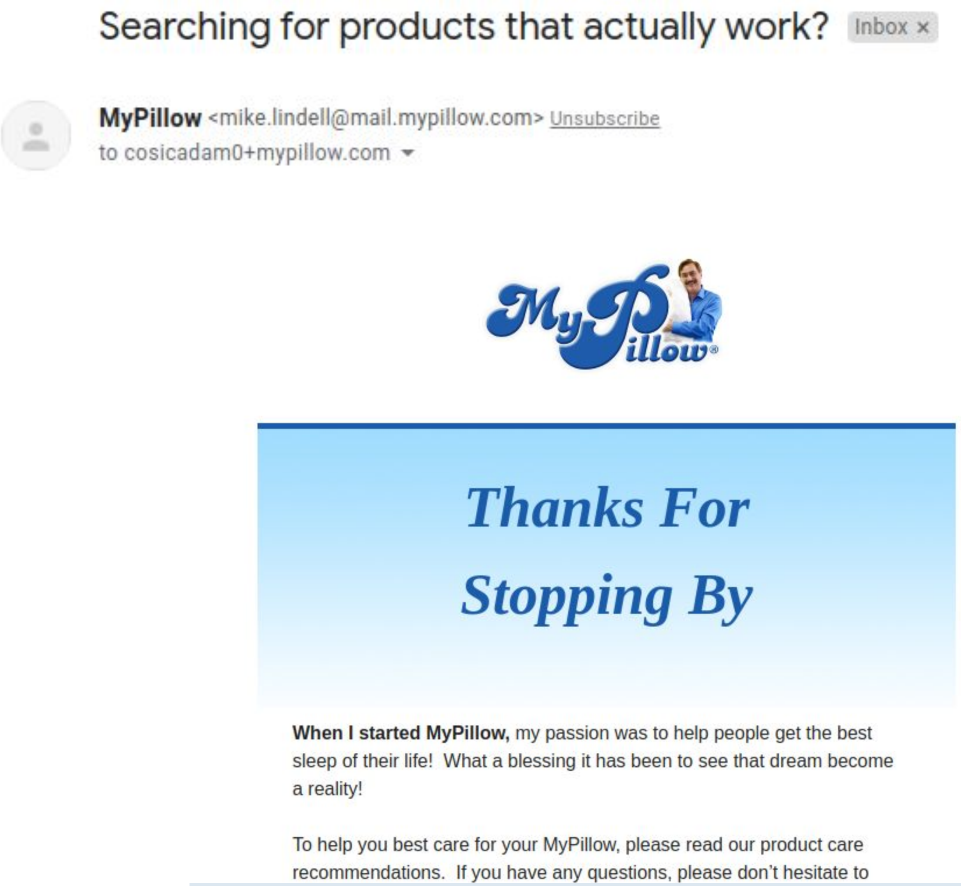
Thank you,

Diabetes UK

*Monday to Friday, 9am to 5pm

View our Privacy and Cookies policies.

Results - Received Emails



Searching for products that actually work? Inbox ×

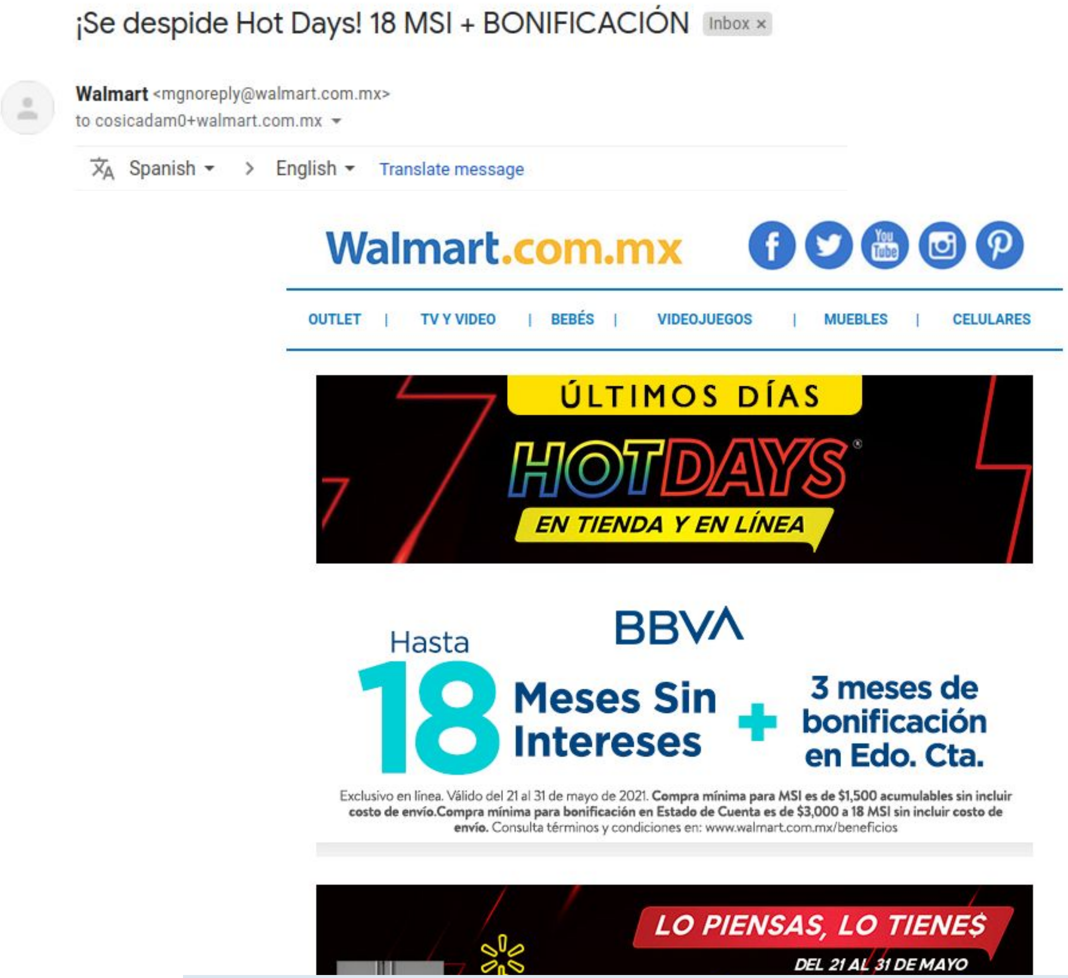
MyPillow <mike.lindell@mail.mypillow.com> Unsubscribe
to cosicadam0+mypillow.com

Thanks For Stopping By

When I started MyPillow, my passion was to help people get the best sleep of their life! What a blessing it has been to see that dream become a reality!

To help you best care for your MyPillow, please read our product care recommendations. If you have any questions, please don't hesitate to

Email from: mypillow.com
Tracker domain: listrakbi.com



¡Se despide Hot Days! 18 MSI + BONIFICACIÓN Inbox ×

Walmart <mgnoreply@walmart.com.mx>
to cosicadam0+walmart.com.mx

Spanish
English
Translate message
Walmart.com.mx
OUTLET | TV Y VIDEO | BEBÉS | VIDEOJUEGOS | MUEBLES | CELULARES
ÚLTIMOS DÍAS
HOTDAYS®
EN TIENDA Y EN LÍNEA
BBVA
Hasta 18 Meses Sin Intereses + 3 meses de bonificación en Edo. Cta.

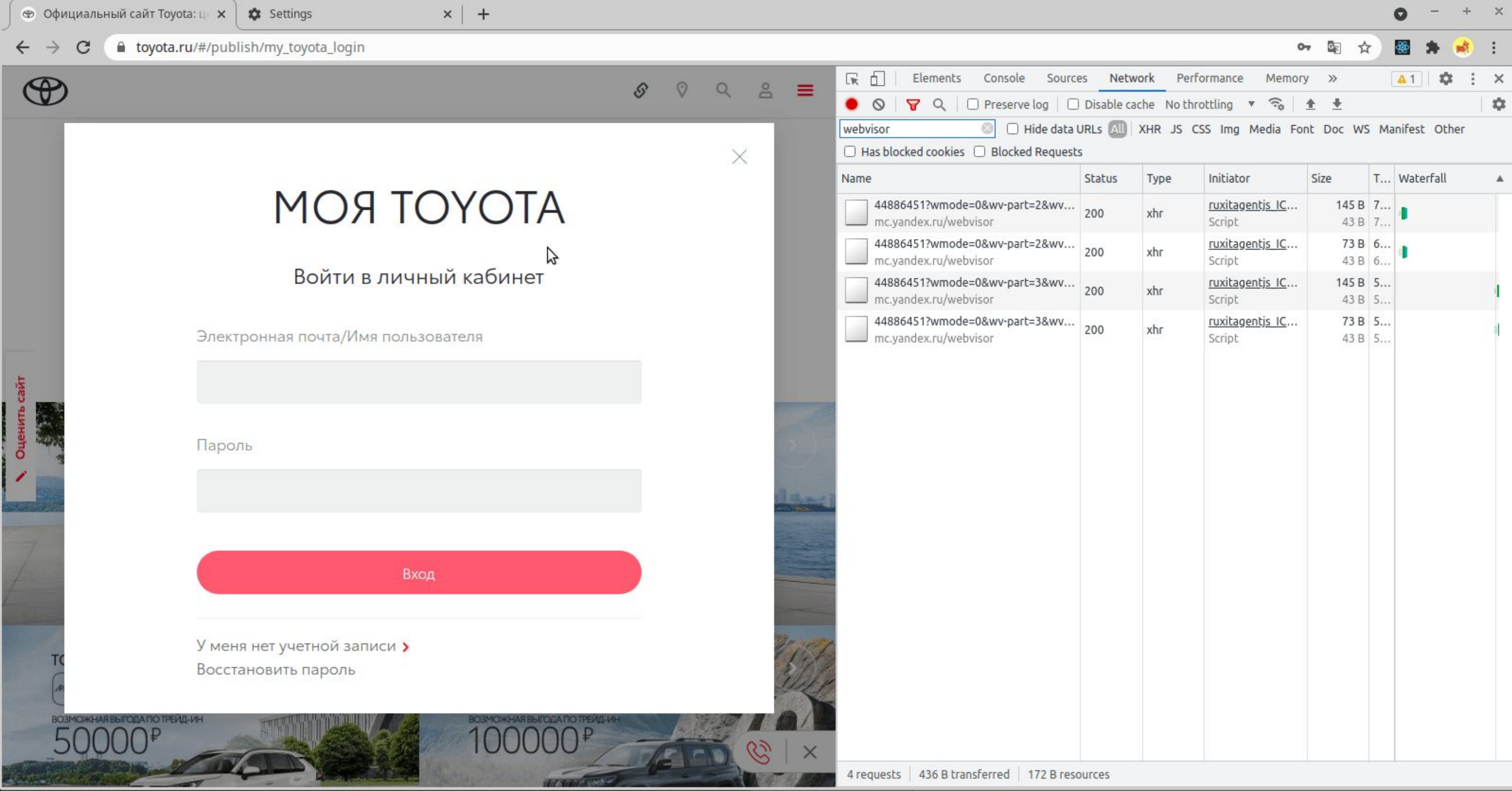
Exclusivo en línea. Válido del 21 al 31 de mayo de 2021. **Compra mínima para MSI es de $1,500 acumulables sin incluir costo de envío.Compra mínima para bonificación en Estado de Cuenta es de $3,000 a 18 MSI sin incluir costo de envío.** Consulta términos y condiciones en: www.walmart.com.mx/beneficios

Email from: walmart.com.mx
Tracker domain: veinteractive.com

Results - Password Leaks on 52 websites

Incidental collection by

- Yandex Metrica: due to React framework (50 websites)
- Mixpanel: due to outdated SDK usage (1 website)
- LogRocket: No response (1 website)



МОЯ TOYOTA

Войти в личный кабинет

Электронная почта/Имя пользователя

Пароль

Вход

У меня нет учетной записи ›

Восстановить пароль

Filter: webvisor

Name	Status	Type	Initiator	Size	T...	Waterfall
44886451?wmode=0&wv-part=2&wv... mc.yandex.ru/webvisor	200	xhr	ruxitagentjs_IC... Script	145 B 43 B	7... 7...	
44886451?wmode=0&wv-part=2&wv... mc.yandex.ru/webvisor	200	xhr	ruxitagentjs_IC... Script	73 B 43 B	6... 6...	
44886451?wmode=0&wv-part=3&wv... mc.yandex.ru/webvisor	200	xhr	ruxitagentjs_IC... Script	145 B 43 B	5... 5...	
44886451?wmode=0&wv-part=3&wv... mc.yandex.ru/webvisor	200	xhr	ruxitagentjs_IC... Script	73 B 43 B	5... 5...	

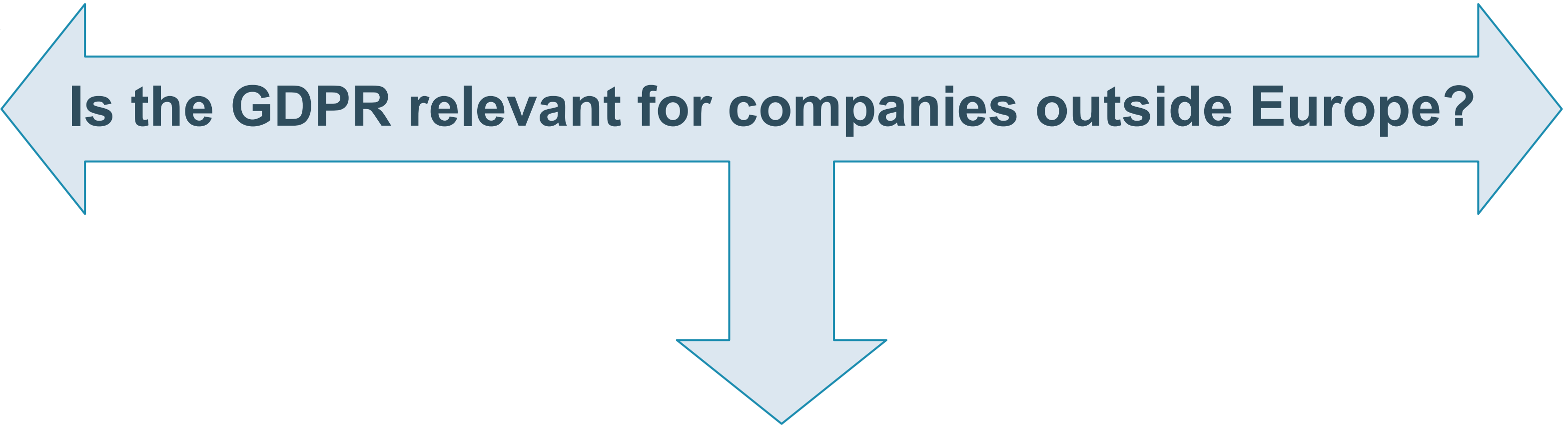
4 requests | 436 B transferred | 172 B resources

Password collection disclosures

- Notified more than 50 websites about the password collection

- Reached out to all third parties
  - Yandex and MixPanel released an update
  - LogRocket never replied
- Reached out to first parties
  - With Russian translation of the email

Does Email Exfiltration Comply With the GDPR?



Is the GDPR relevant for companies outside Europe?

Is the controller is based in the EU?

Do the company 'monitors' the behavior of people in the EU?

Does the company offers goods or services to Europeans?

Does Email Exfiltration Comply With the GDPR?

1 **Transparency principle**

Personal data must be processed ‘fairly and in a transparent manner’

2 **Purpose limitation principle**

Controllers can only collect personal data if they specify a clear purpose in advance

3 **The requirement for a legal basis such as consent**

The controller always needs a ‘legal basis’ to process personal data

GDPR Requests

52%

30/58 first parties replied

- Were not aware & removed
  - fivethirtyeight.com (via Walt Disney's DPO)
  - trello.com (Atlassian)
  - lever.co, branch.io and cision.com
- Marriott: Glassbox is used for **customer care, technical support, and fraud prevention**
- Tapad, not offering their services to UK & EEA users since August, 2021

GDPR Requests

53%

15/28 third parties replied

- Adobe and Yandex: Referred to corresponding first parties
- Taboola: CMP misconfiguration
- Adroll: Promised to rollout a comprehensive solution

Outreach to identified websites in the US crawl

0%

0/33 first parties replied

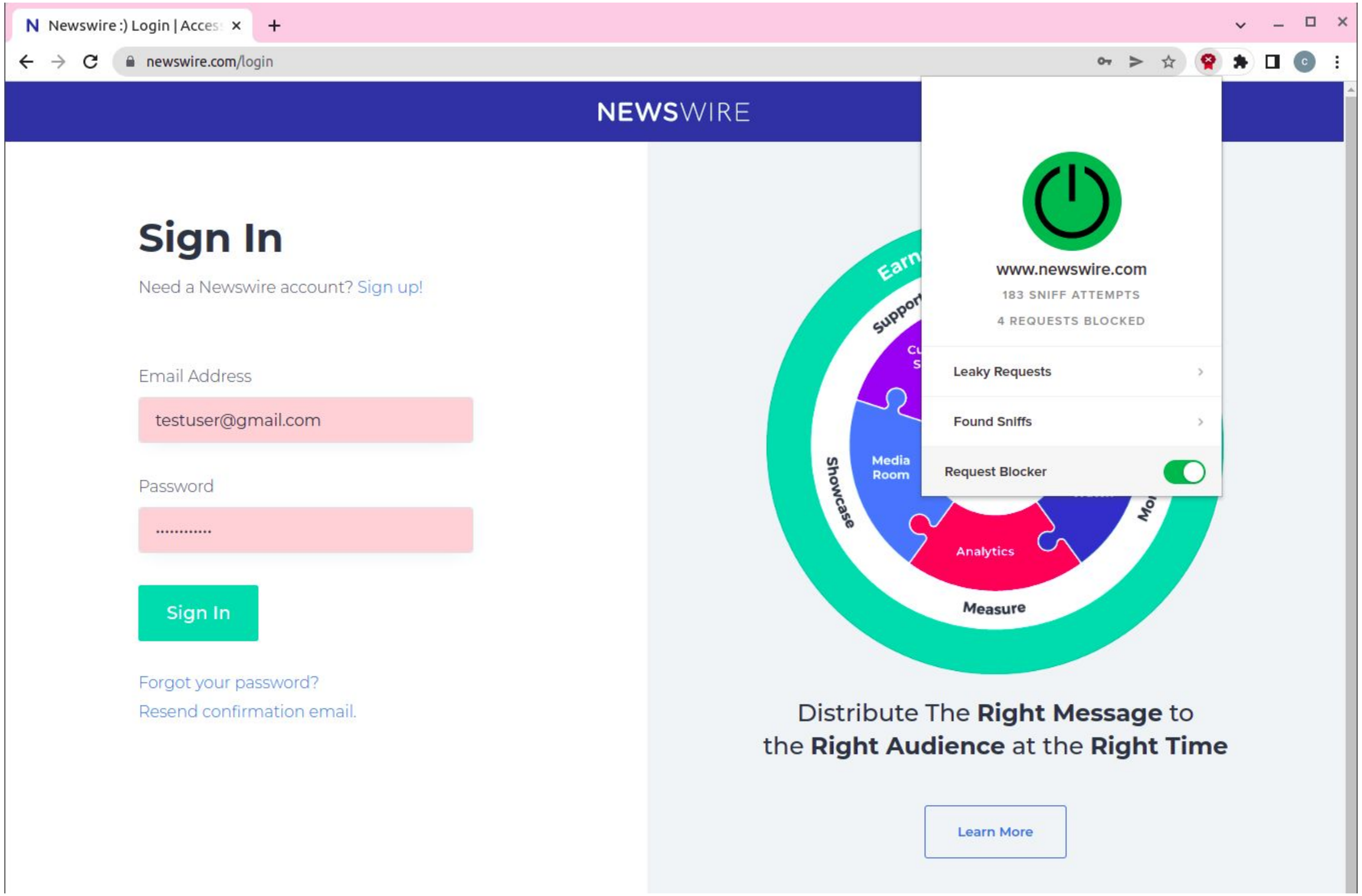
- We sent a friendly notification rather than a formal GDPR request.
- We did not get any response from these 33 websites.

Countermeasures

- Browser add-ons **block requests** to tracker domains
- Private **email relay** services hide users’ emails
  - Apple, Mozilla, DuckDuckGo
  - e.g. testuser@duck.com-> testuser@gmail.com
- **NO** tool for detection and prevention of **sniff & exfiltration** on online forms

Browser add-on: LEAKINSPECTOR

- Detect sniff attempts
- Request Blocker
  - Filters & blocks leaky requests to trackers
- The UI is based on DuckDuckGo Privacy Essentials



Sign In

Need a Newswire account? Sign up!

Email Address: testuser@gmail.com

Password: ....

Forgot your password?

Resend confirmation email.

www.newswire.com

183 SNIFF ATTEMPTS

4 REQUESTS BLOCKED

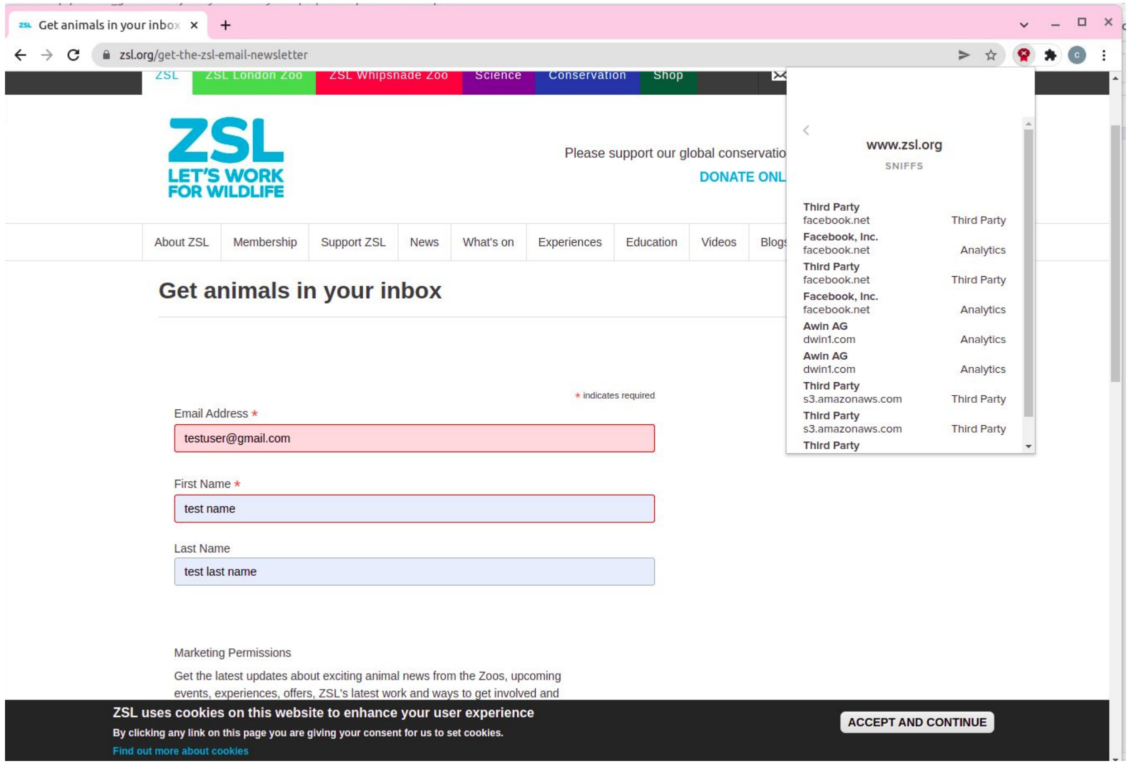
Request Blocker: on

Distribute The **Right Message** to the **Right Audience** at the **Right Time**

Implementation Details - Detection of Sniff Attempts

Check all interesting input fields

- Modify getter method of HTMLInputElement
- Process the stack trace
- Compose a list of script origins
- Check the script domain is same or not with the domain of the page URL
- Highlight the sniffed input field



Please support our global conservatio

Get animals in your inbox

* indicates required

Email Address *
testuser@gmail.com

First Name *
test name

Last Name
test last name

Marketing Permissions

Get the latest updates about exciting animal news from the Zoos, upcoming events, experiences, offers, ZSL's latest work and ways to get involved and

ZSL uses cookies on this website to enhance your user experience

By clicking any link on this page you are giving your consent for us to set cookies.

Find out more about cookies

www.zsl.org

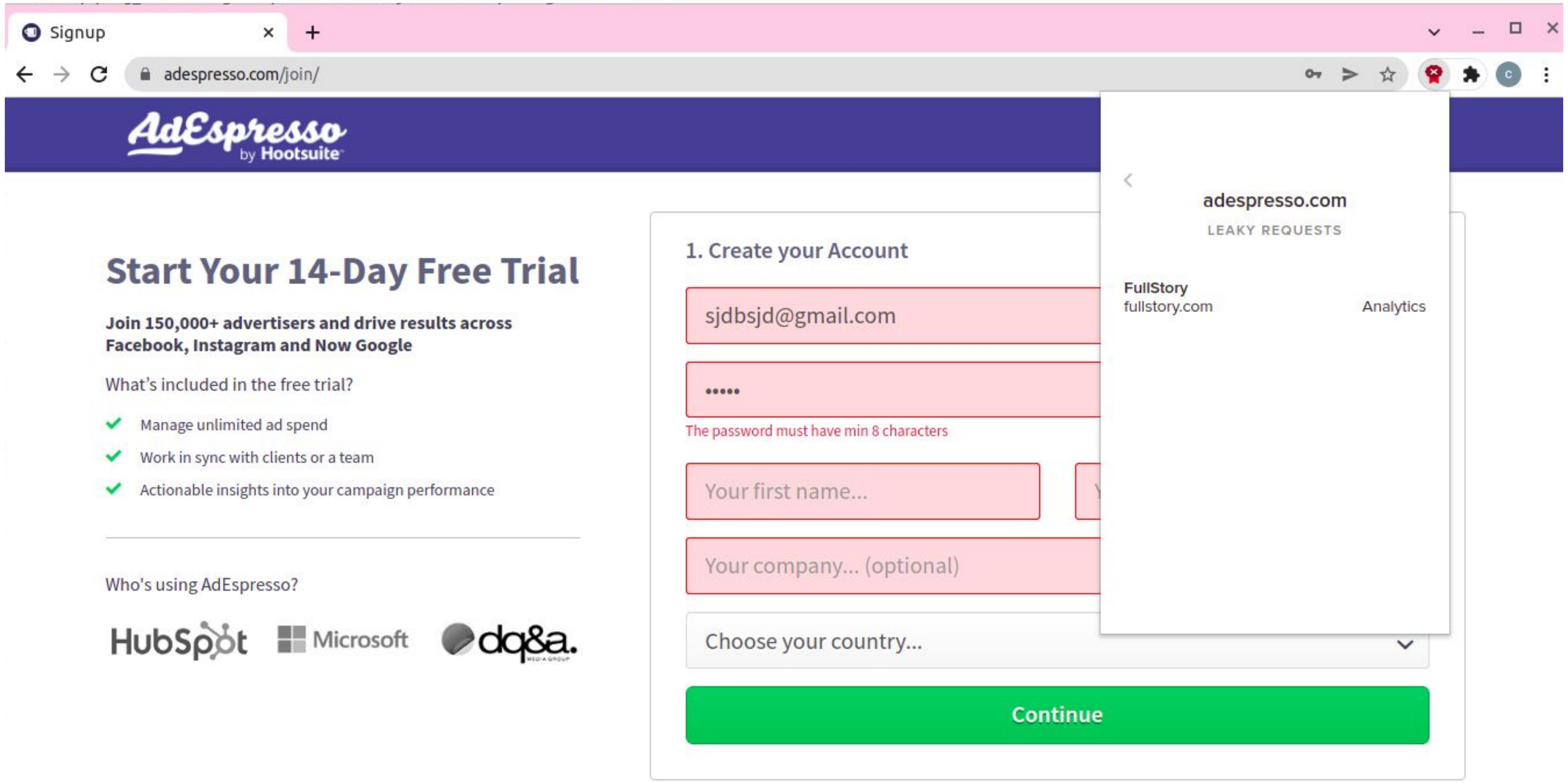
SNIFFS

Third Party facebook.net	Third Party
Facebook, Inc. facebook.net	Analytics
Third Party facebook.net	Third Party
Facebook, Inc. facebook.net	Analytics
Awin AG dwin1.com	Analytics
Awin AG dwin1.com	Analytics
Third Party s3.amazonaws.com	Third Party
Third Party s3.amazonaws.com	Third Party
Third Party	

Implementation Details- Detection of Leaky Requests

Before each request was sent to the its end points

- Check the POST body and URL in the request
- Encoded, hashed, obfuscated, cleartext version of the input value is in the request
- Leak detector implemented based on methodology proposed by Englehardt et al. (PETS’18) [3]



Signup

adespresso.com/join/

AdEspresso by Hootsuite

Start Your 14-Day Free Trial

Join 150,000+ advertisers and drive results across Facebook, Instagram and Now Google

What’s included in the free trial?

- ✓ Manage unlimited ad spend
- ✓ Work in sync with clients or a team
- ✓ Actionable insights into your campaign performance

Who’s using AdEspresso?

HubSpot Microsoft dq&a.

1. Create your Account

sjdbsjd@gmail.com

•••••

The password must have min 8 characters

Your first name...

Your company... (optional)

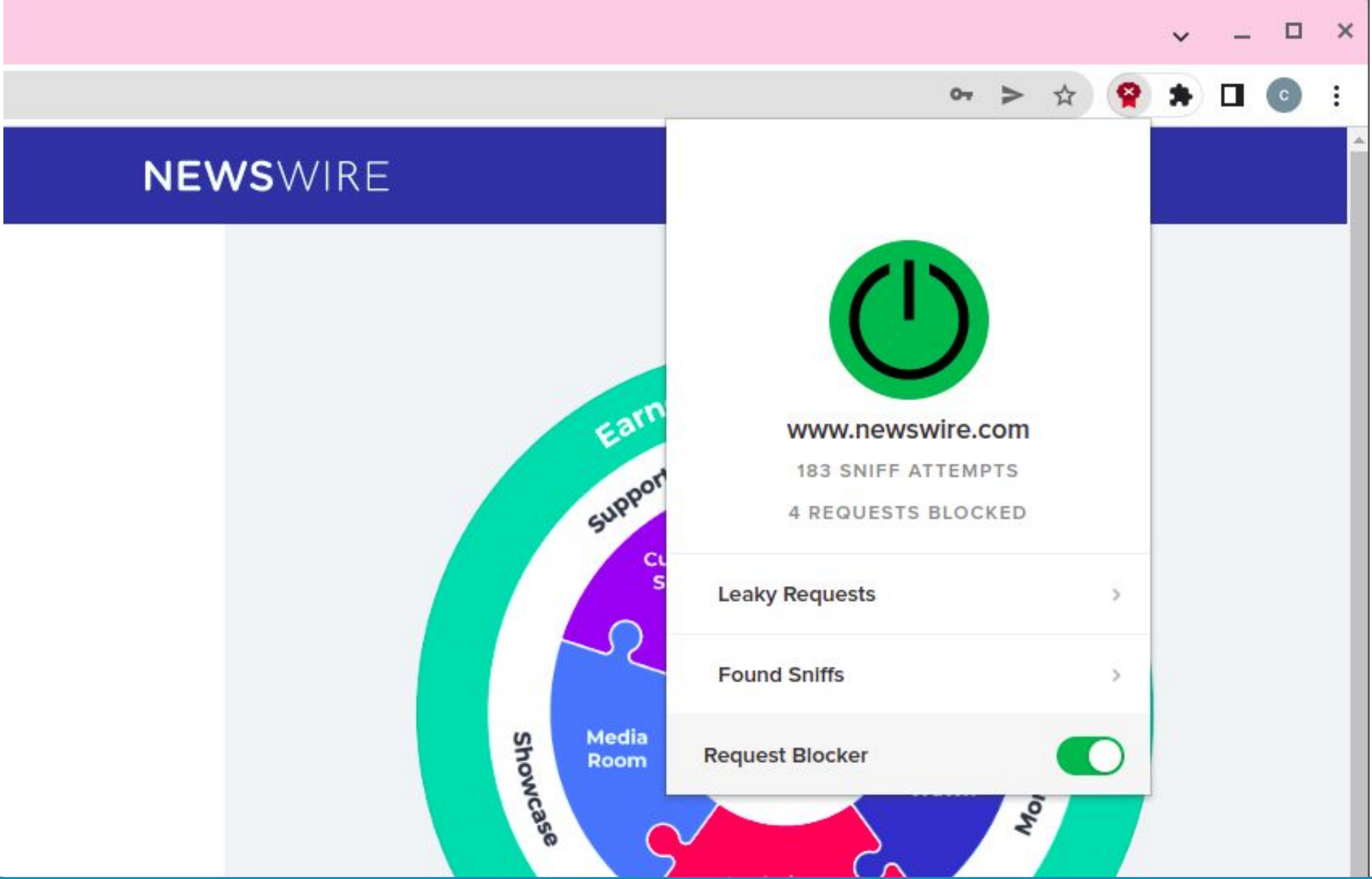
Choose your country...

Continue

adespresso.com

LEAKY REQUESTS

FullStory fullstory.com	Analytics

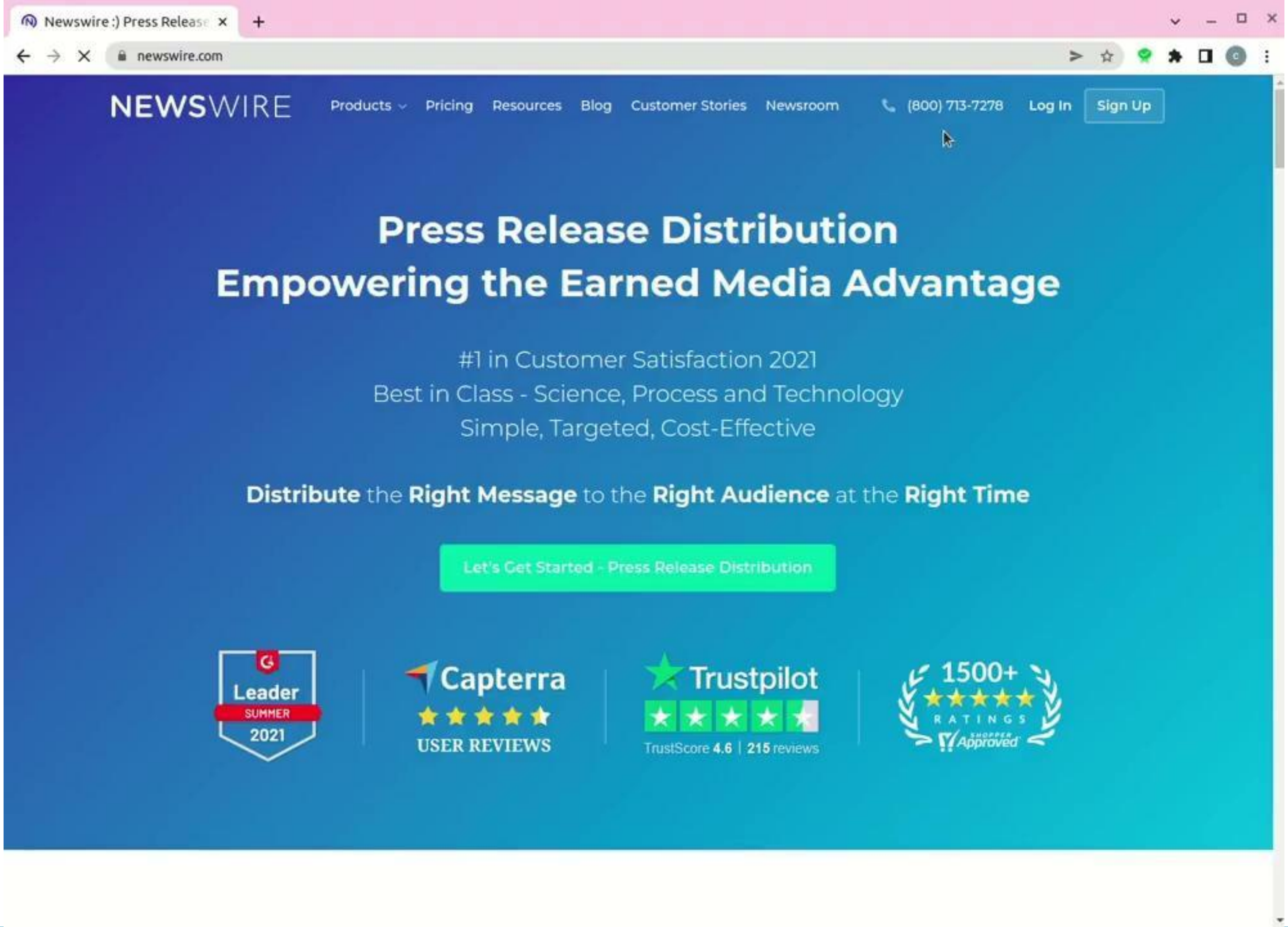


www.newswire.com

183 SNIFF ATTEMPTS

4 REQUESTS BLOCKED

Request Blocker: on



Press Release Distribution
Empowering the Earned Media Advantage

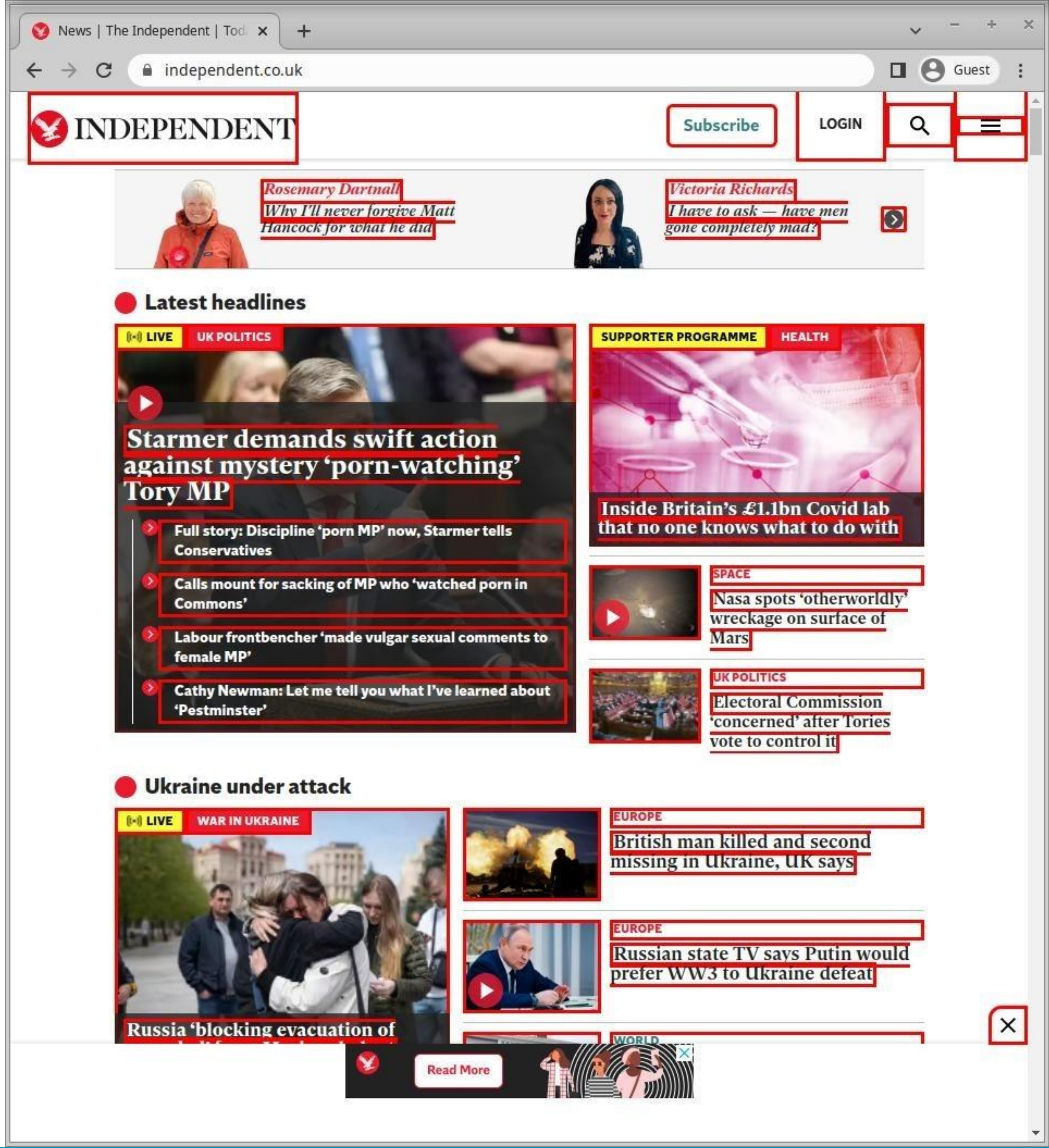
#1 in Customer Satisfaction 2021
Best in Class - Science, Process and Technology
Simple, Targeted, Cost-Effective

Distribute the **Right Message** to the **Right Audience** at the **Right Time**

Let's Get Started - Press Release Distribution
Leader
SUMMER
2021
Capterra
USER REVIEWS
Trustpilot
TrustScore 4.6 | 215 reviews
1500+
RATINGS
Shopper Approved

Leaks to Facebook & TikTok

- Triggered when the user clicks any button on the page.



Rosemary Dartnall
Why I'll never forgive Matt Hancock for what he did

Victoria Richards
I have to ask — have men gone completely mad?

Latest headlines

LIVE UK POLITICS

Starmer demands swift action against mystery 'porn-watching' Tory MP

Full story: Discipline 'porn MP' now, Starmer tells Conservatives

Calls mount for sacking of MP who 'watched porn in Commons'

Labour frontbencher 'made vulgar sexual comments to female MP'

Cathy Newman: Let me tell you what I've learned about 'Pestminster'

SUPPORTER PROGRAMME HEALTH

Inside Britain's £1.1bn Covid lab that no one knows what to do with

SPACE

Nasa spots 'otherworldly' wreckage on surface of Mars

UK POLITICS

Electoral Commission 'concerned' after Tories vote to control it

Ukraine under attack

LIVE WAR IN UKRAINE

Russia 'blocking evacuation of

EUROPE

British man killed and second missing in Ukraine, UK says

EUROPE

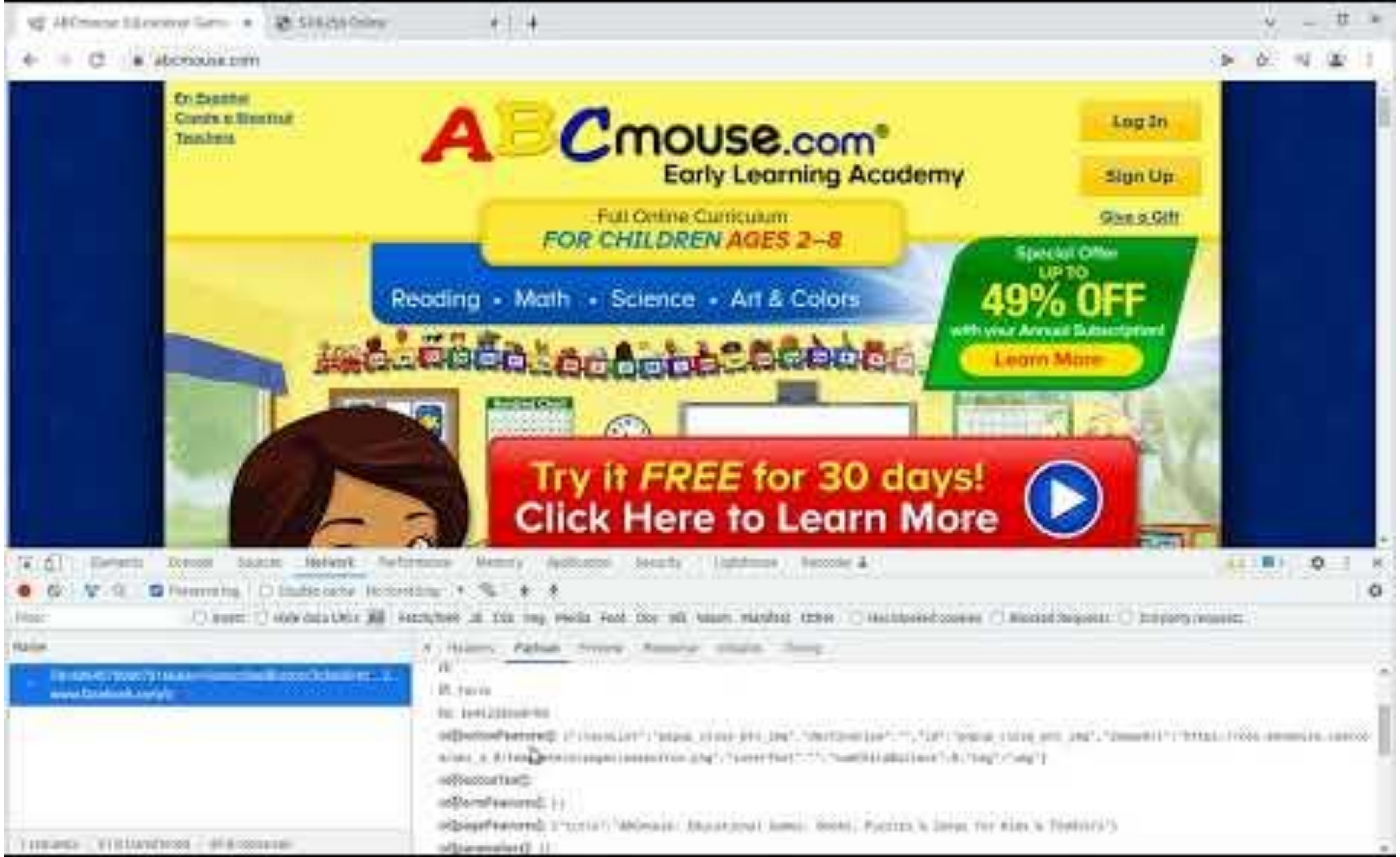
Russian state TV says Putin would prefer WW3 to Ukraine defeat

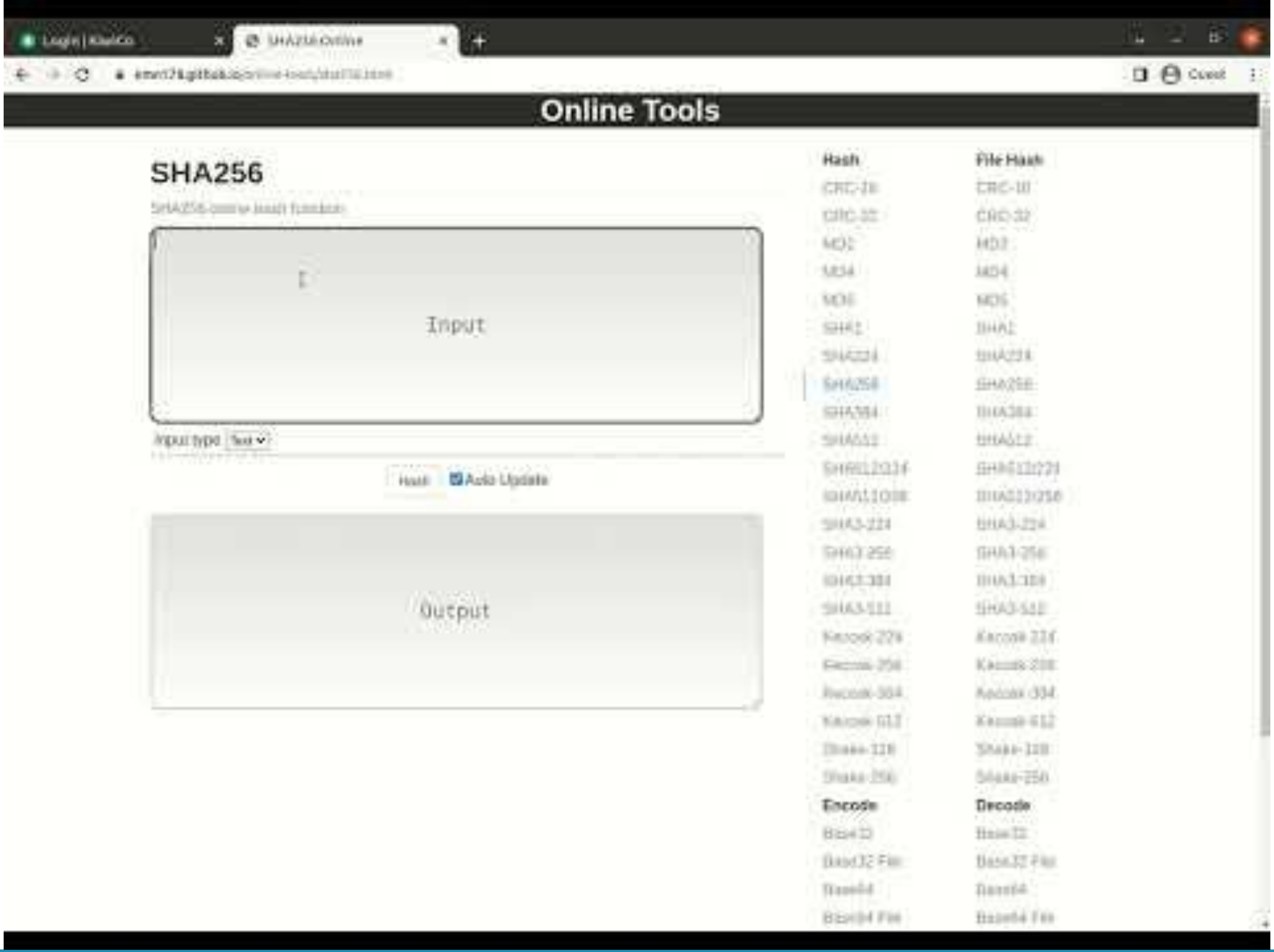
WORLD

Leaks to Facebook & TikTok

- Disclosed to Facebook and TikTok

	EU	US
Facebook	7,379	8,438
TikTok	147	154


abcmouse.com
En Español
Create a Shortcut
Teachers
ABCmouse.com®
Early Learning Academy
Log In
Sign Up
Give a Gift
Full Online Curriculum
FOR CHILDREN AGES 2–8
Reading • Math • Science • Art & Colors
Special Offer
UP TO
49% OFF
with your Annual Subscription!
Learn More
Try it FREE for 30 days!
Click Here to Learn More
Network
Payload


Login | KiwiCo
SHA256 Online
emn178.github.io/online-tools/sha256.html
Guest
Online Tools
SHA256
SHA256 online hash function
Input
Input type Text
Hash
☑ Auto Update
Output
Hash
CRC-16
CRC-32
MD2
MD4
MD5
SHA1
SHA224
SHA256
SHA384
SHA512
SHA512/224
SHA512/256
SHA3-224
SHA3-256
SHA3-384
SHA3-512
Keccak-224
Keccak-256
Keccak-384
Keccak-512
Shake-128
Shake-256
File Hash
CRC-16
CRC-32
MD2
MD4
MD5
SHA1
SHA224
SHA256
SHA384
SHA512
SHA512/224
SHA512/256
SHA3-224
SHA3-256
SHA3-384
SHA3-512
Keccak-224
Keccak-256
Keccak-384
Keccak-512
Shake-128
Shake-256
Encode
Base32
Base32 File
Base64
Base64 File
Decode
Base32
Base32 File
Base64
Base64 File

Wrap up

- Email leaks on 1844, 2950 websites, EU, US, resp.
- Password leaks on 52 websites due to session replay scripts
- Uncovered 41 unlisted tracking domains
- Developed a transparency extension that detects PII sniffs and leaks

Any Questions?

References

[1] E. Papadogiannakis, P. Papadopoulos, N. Kourtellis, and E. P. Markatos. “User Tracking in the Postcookie Era: How Websites Bypass GDPR Consent to Track Users”. In: Proceedings of the Web Conference 2021. 2021, pp. 2130–2141.

[2] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal. “Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence”.In: Proceedings of the 2020 CHI conference on human factors in computing systems. 2020, pp. 1–13.

[3] S. Englehardt, J. Han, and A. Narayanan. “I never signed up for this! Privacy implications of email tracking”. In: Proc. Priv. Enhancing Technol. 2018.1 (2018), pp. 109–126.

[4] Oleksii Starov, Phillipa Gill, and Nick Nikiforakis. FormLock. https://github.com/ostarov/Formlock, 2021.

[5] Surya Mattu and Kashmir Hill. Before You Hit ’Submit,’ This Company Has Already Logged Your Personal Data. Gizmodo, 2017.https://gizmodo.com/before-you-hit-submit-this-company-has-already-logge-1795906081

[6] 6 Steps for Avoiding Online Form Abandonment. https://themanifest.com/web-design/blog/6-steps-avoid-online-form-abandonment, Feb 2022. [Online; accessed 2022-02-12].

[7] Rolf Bagge, Célestin Matte, Éric Daspet, Kaspar Emanuel, Sam Macbeth, and Steven Roeland.Consent-O-Matic.https://github.com/cavi-au/Consent-O-Matic/, 2019.

[8] Mozilla Fathom documentation. https://mozilla. github.io/fathom/.[Online; accessed 2021-06-01].