Midas Nouwens

Assistant Professor
Digital Harms & Rights

@midasnouwens

TRACKING TECHNOLOGIES

DATA PROTECTION REGULATION

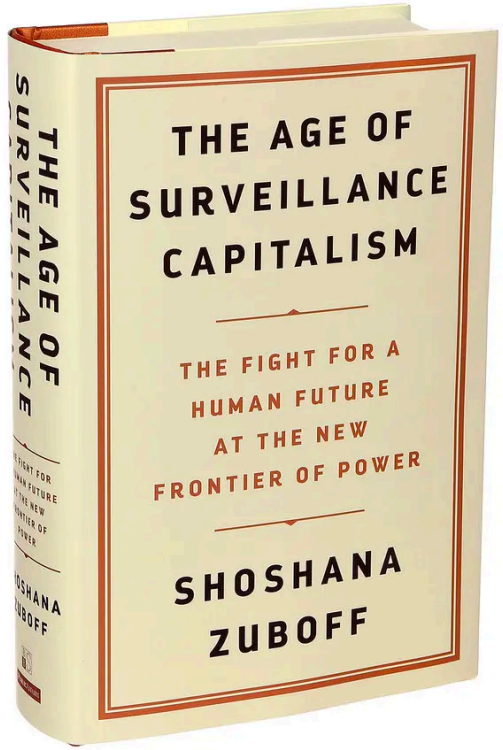
REGULATORY COMPLIANCE

MAIN TRACKING TECHNOLOGIES (WEB, MOBILE, IOT)

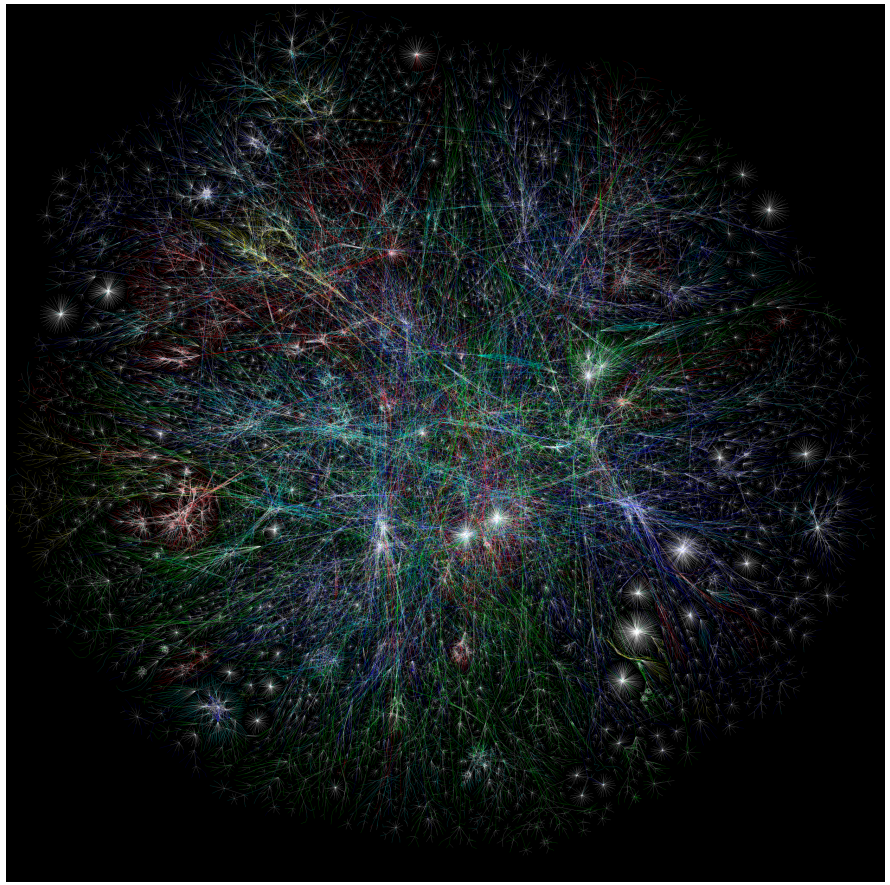
Tracking is the workhorse of surveillance capitalism. Without it, the vast wealth and power of large digital platforms would not be possible.

- IP address
- Cookies
- Fingerprinting
- Tracking pixels
- Software Development Kits ('plug-and-play' services)
- (Coming soon: FLoC, Federated Learning of Cohorts)

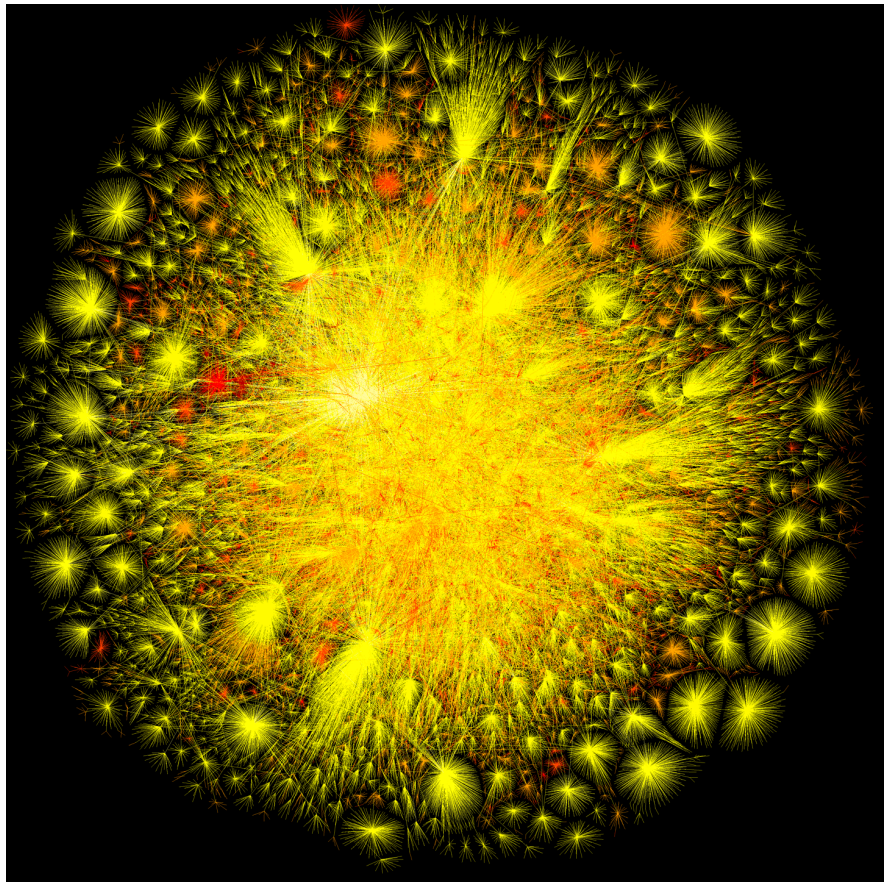
identification + segmentation => targeting



IP ADDRESS

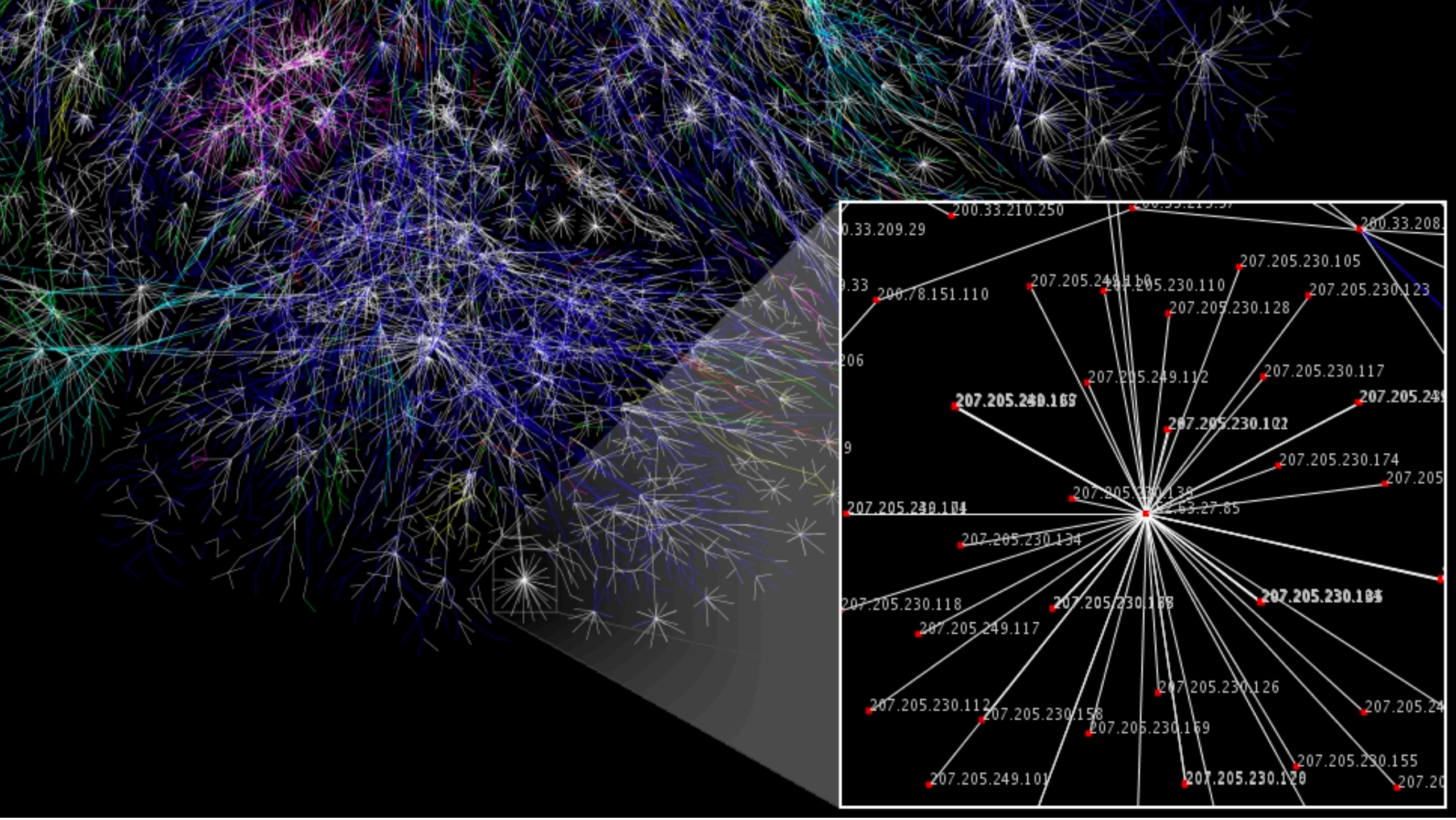


The internet in 2003



The internet in 2010

https://www.opte.org/the-internet

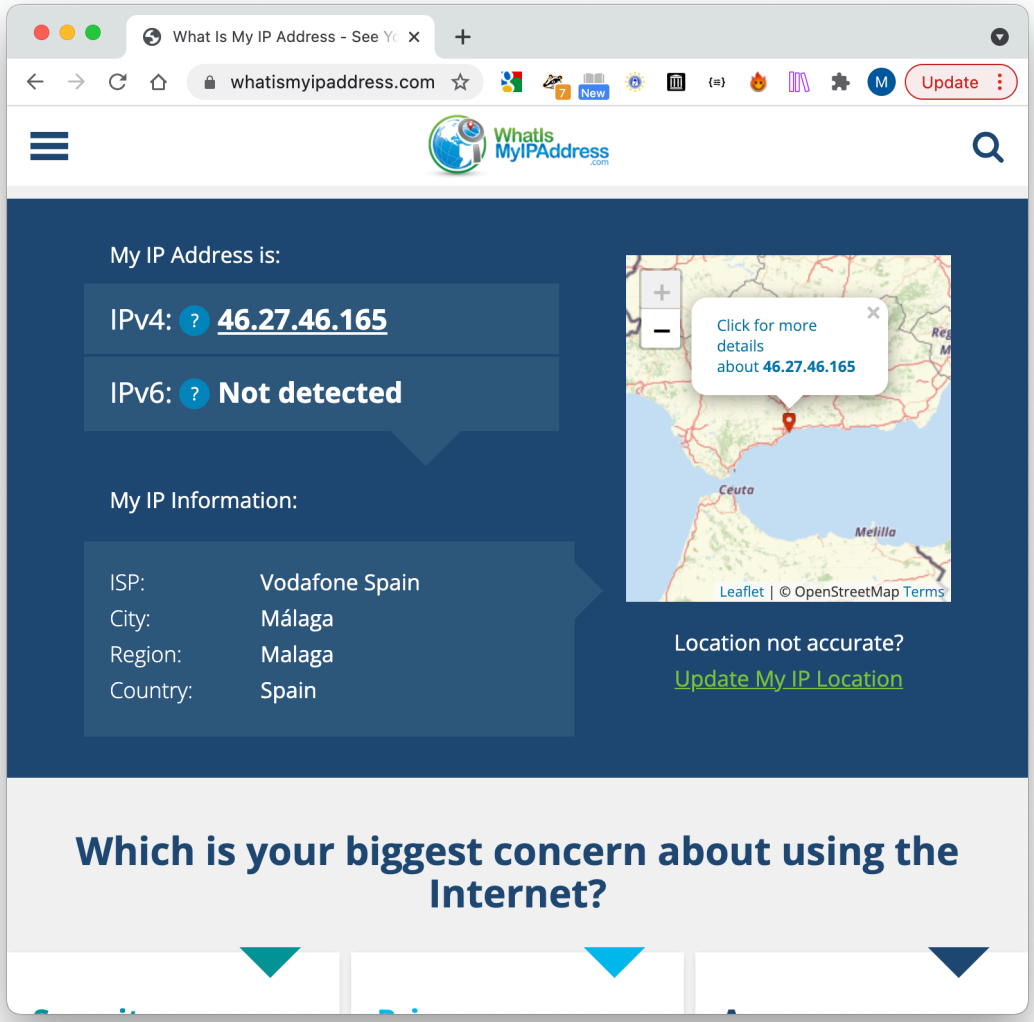

200.33.210.250
207.205.230.105
207.205.230.110
207.205.230.123
207.205.230.128
207.205.249.112
207.205.230.117
207.205.230.174
207.205.230.134
207.205.230.118
207.205.249.117
207.205.230.126
207.205.230.112
207.205.230.158
207.205.230.169
207.205.230.155
207.205.249.101

IP ADDRESS

85%+ of users retain the same IP address for more than a month

70% of users retain it for more than 2 months

2% of users retain it for more than 100 days



Websites can use it to identify:

- where in the world visitors are coming from
- a particular user, so they can track
  - how many times you visit their website
  - how long you stay on their website

Mishra, V., Laperdrix, P., Vastel, A., Rudametkin, W., Rouvoy, R., Lopatka, M.: Don't count me out: On the relevance of IP address in the tracking ecosystem.

COOKIES

A file with data that is saved and read by a website on your computer

Can contain any information

Invented in 1995 to save information from a user between website visits

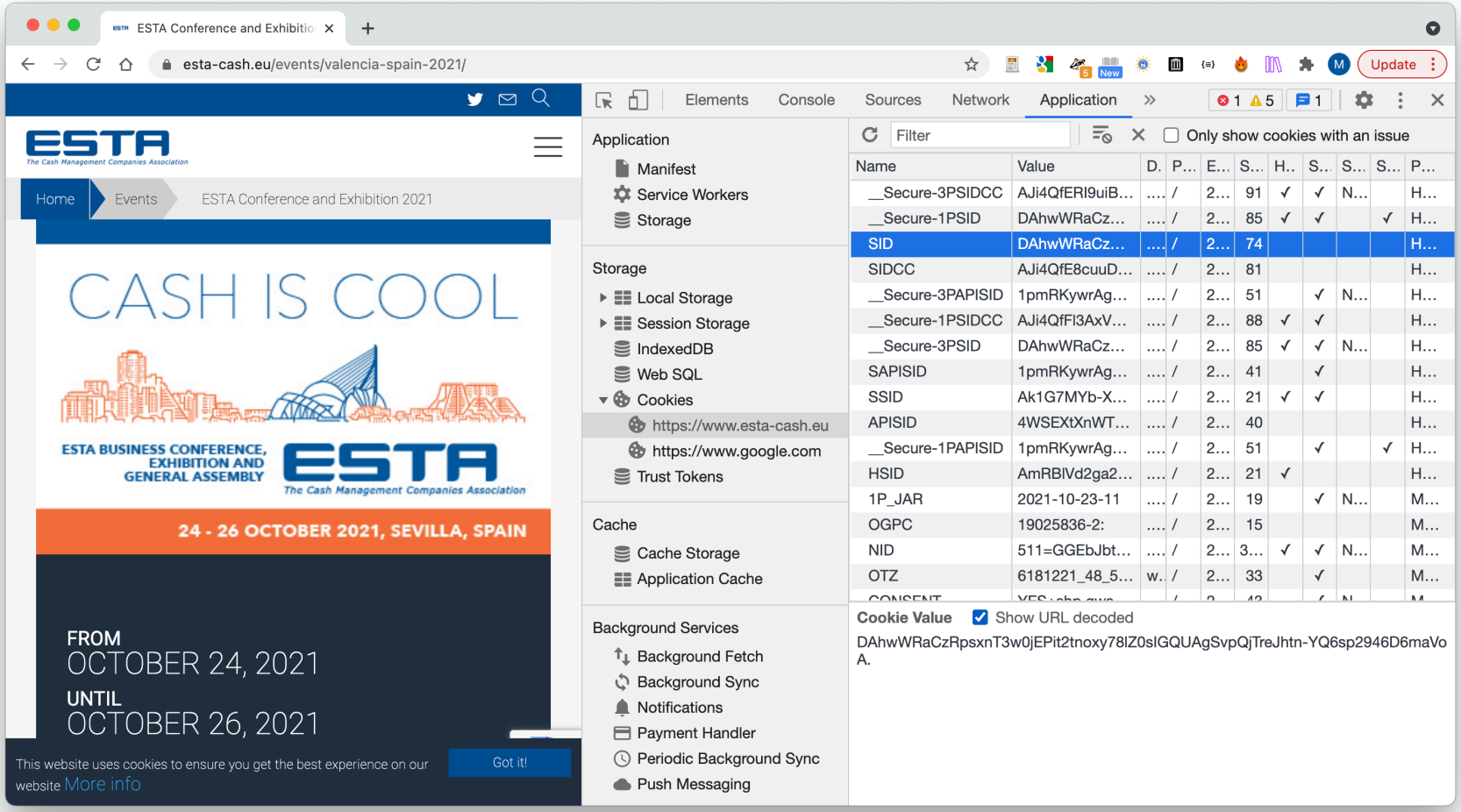
- shopping cart
- language settings

```
Syntax of the Set-Cookie HTTP Response Header:
Set-Cookie: NAME=OPAQUE\_STRING \
    [; expires= ] \
    [; path=] \
    [; domain=] \
    [; secure]

Syntax of the Cookie HTTP Request Header:
Cookie: NAME=OPAQUE\_STRING *[; NAME=OPAQUE\_STRING]
```

Original proposal for the cookie specifications made to a mailing list by Lou Montulli in 1995

COOKIES



Right click > Inspect > Application > Cookies

COOKIES

First party cookies:

Cookies that are placed by the website domain you are visiting

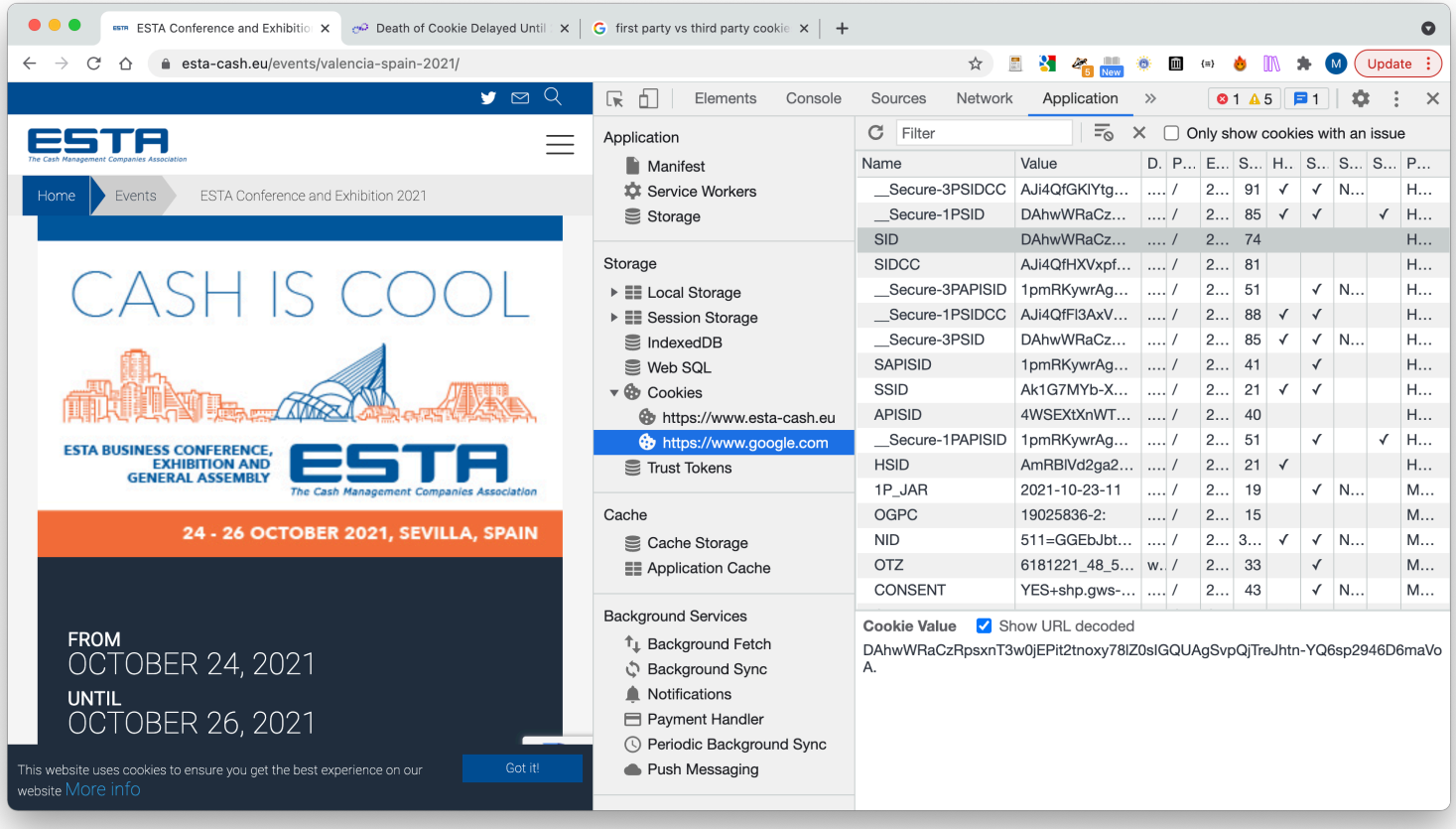
Third party cookies:

Cookies that are placed by other services that are used by the website you are visiting.

Only the domain (e.g., esta-cash.eu) that created the cookie can read and edit the content

With the transition from web 1.0 to web 2.0, third-parties started to play a more important role in tracking. If multiple websites use the same third-party service, this service can use cookies to track your cross-website behaviour

COOKIES

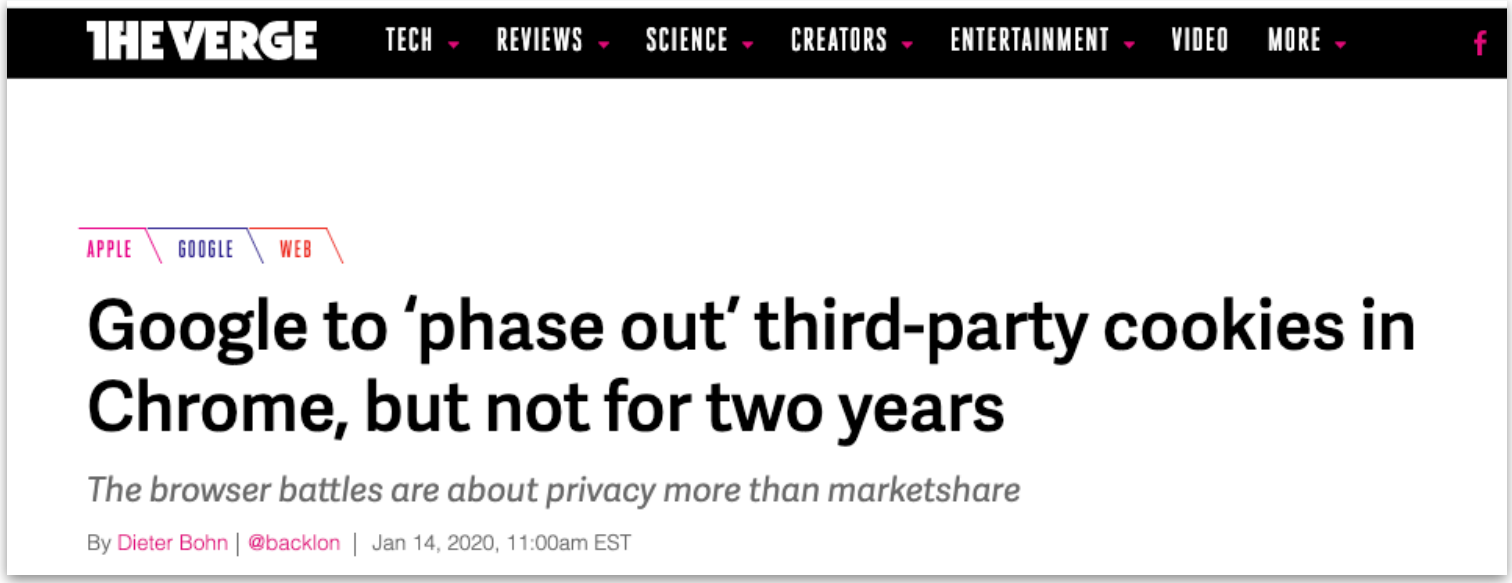


Some cookies are used to authenticate users, helping ensure that only the actual owner of an account can access that account. For example, cookies called 'SID' and 'HSID' contain digitally signed and encrypted records of a user's Google Account ID and most recent sign-in time. The combination of these cookies allows us to block many types of attack, such as attempts to steal the content of forms submitted in Google services.

https://policies.google.com/technologies/cookies?hl=en-US

Today's Firefox Blocks Third-Party Tracking Cookies and Cryptomining by Default

SEPTEMBER 3, 2019 MARISSA WOOD


THE VERGE TECH REVIEWS SCIENCE CREATORS ENTERTAINMENT VIDEO MORE

APPLE GOOGLE WEB

Google to 'phase out' third-party cookies in Chrome, but not for two years

The browser battles are about privacy more than marketshare

By Dieter Bohn | @backlon | Jan 14, 2020, 11:00am EST

https://blog.mozilla.org/en/products/firefox/todays-firefox-blocks-third-party-tracking-cookies-and-cryptomining-by-default/
https://www.theverge.com/2020/1/14/21064698/google-third-party-cookies-chrome-two-years-privacy-safari-firefox

FINGERPRINTING

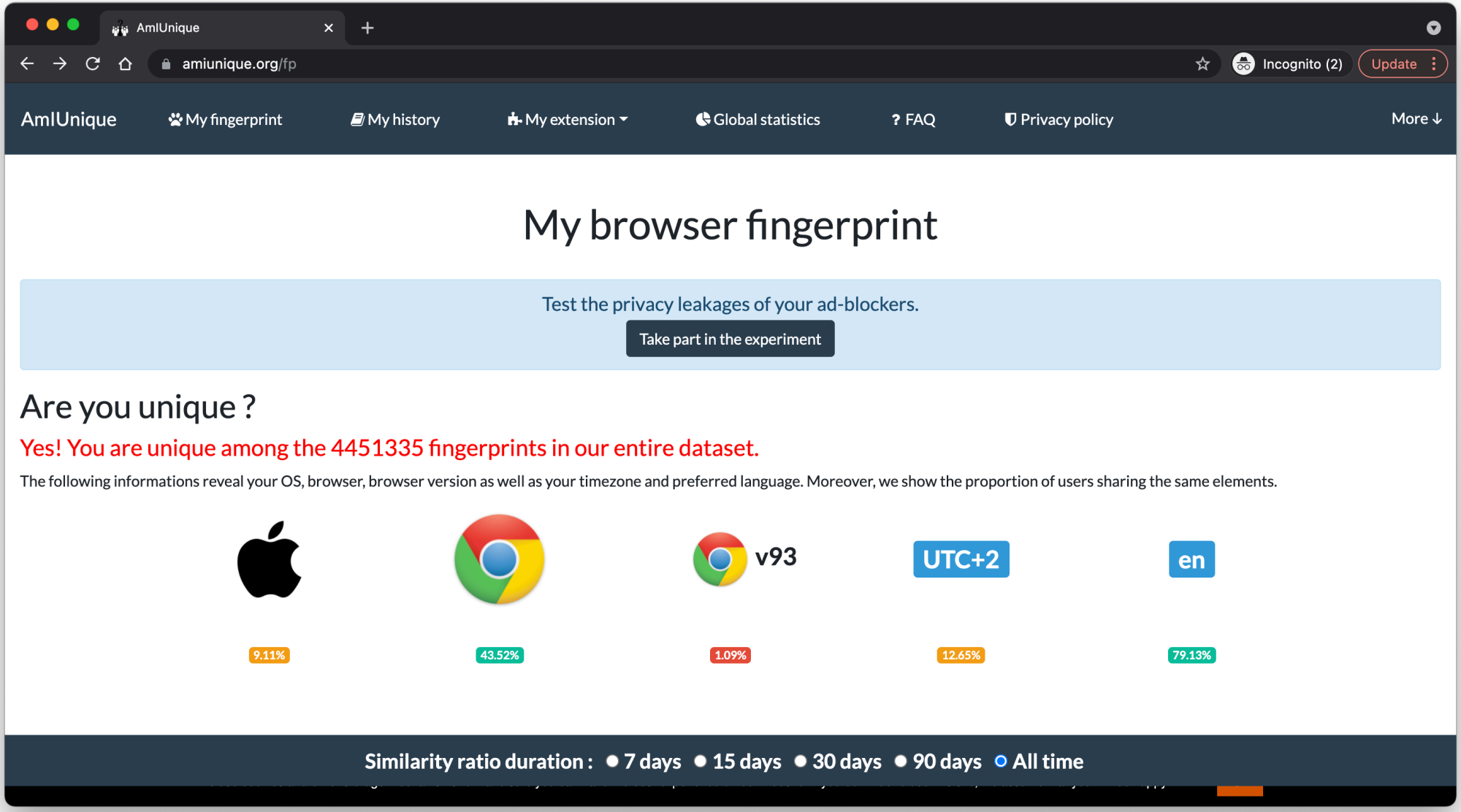
Data collected to (probabilistically) identify a specific device

Data includes device model, screen resolution, OS, preferred language, fonts installed, plugins or extensions installed, how you process graphical content, etc

Powerful because of the combination of all this hard to change information: triangulation + stable

Table 1. Example of a browser fingerprint.

Attribute	Source	Example
User agent	HTTP header	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/64.0.3282.119 Safari/537.36
Accept	HTTP header	text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8
Content encoding	HTTP header	gzip, deflate, br
Content language	HTTP header	en-US,en;q=0.9
List of plugins	JavaScript	Plugin 1: Chrome PDF Plugin. Plugin 2: Chrome PDF Viewer. Plugin 3: Native Client. Plugin 4: Shockwave Flash...
Cookies enabled	JavaScript	yes
Use of local/session storage	JavaScript	yes
Timezone	JavaScript	-60 (UTC+1)
Screen resolution and color depth	JavaScript	1920x1200x24
List of fonts	Flash or JS	Abyssinica SIL,Aharoni CLM,AR PL UMing CN,AR PL UMing HK,AR PL UMing TW...
List of HTTP headers	HTTP headers	Referer X-Forwarded-For Connection Accept Cookie Accept-Language Accept-Encoding User-Agent Host
Platform	JavaScript	Linux x86_64
Do Not Track	JavaScript	yes
Canvas	JavaScript	Cwm fjordbank glyphs vext quiz, 😃 Cwm fjordbank glyphs vext quiz, 😃
WebGL Vendor	JavaScript	NVIDIA Corporation
WebGL Renderer	JavaScript	GeForce GTX 650 Ti/PCIe/SSE2
Use of an ad blocker	JavaScript	yes



https://amiunique.org/

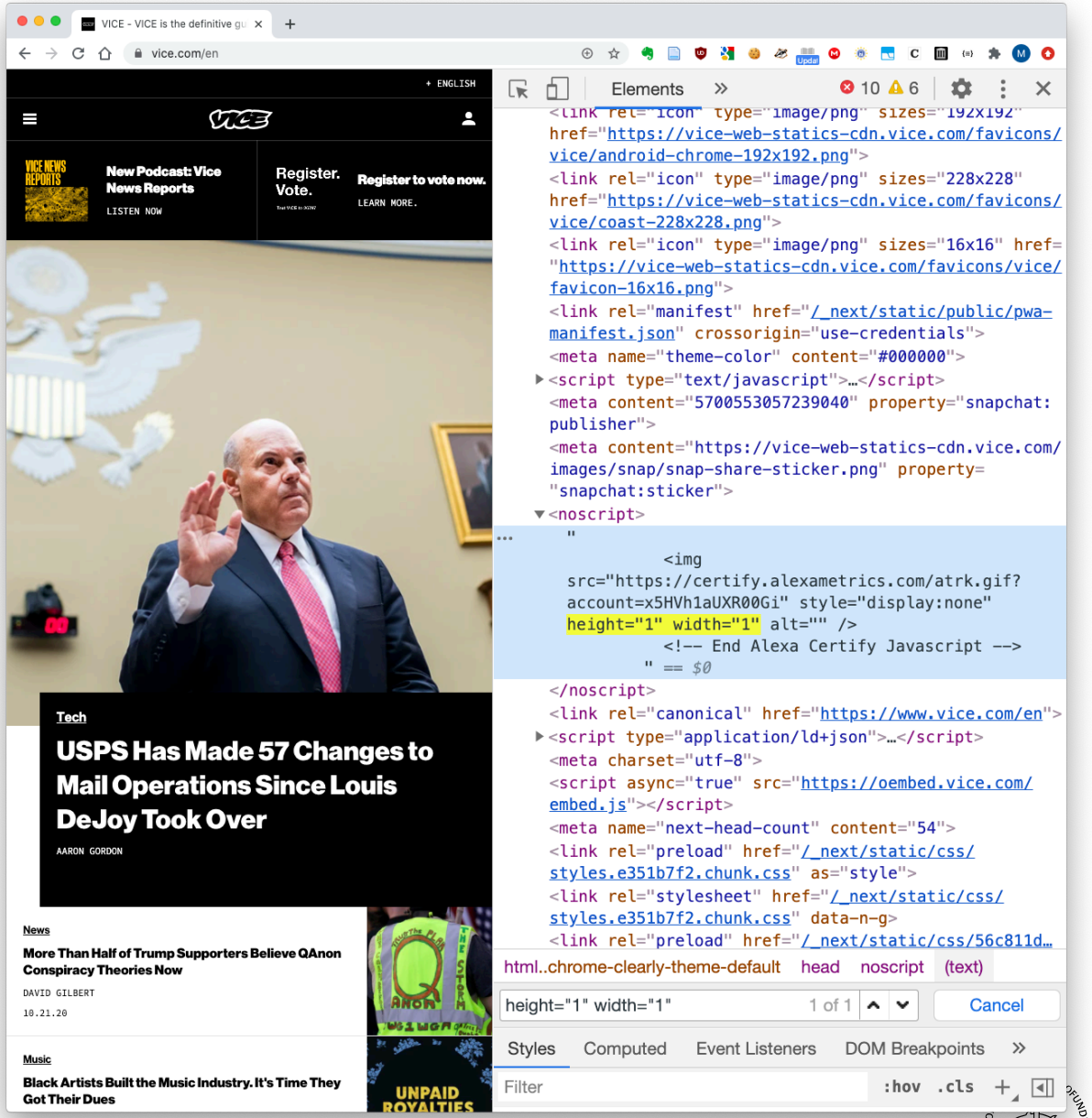
PIXEL TRACKING

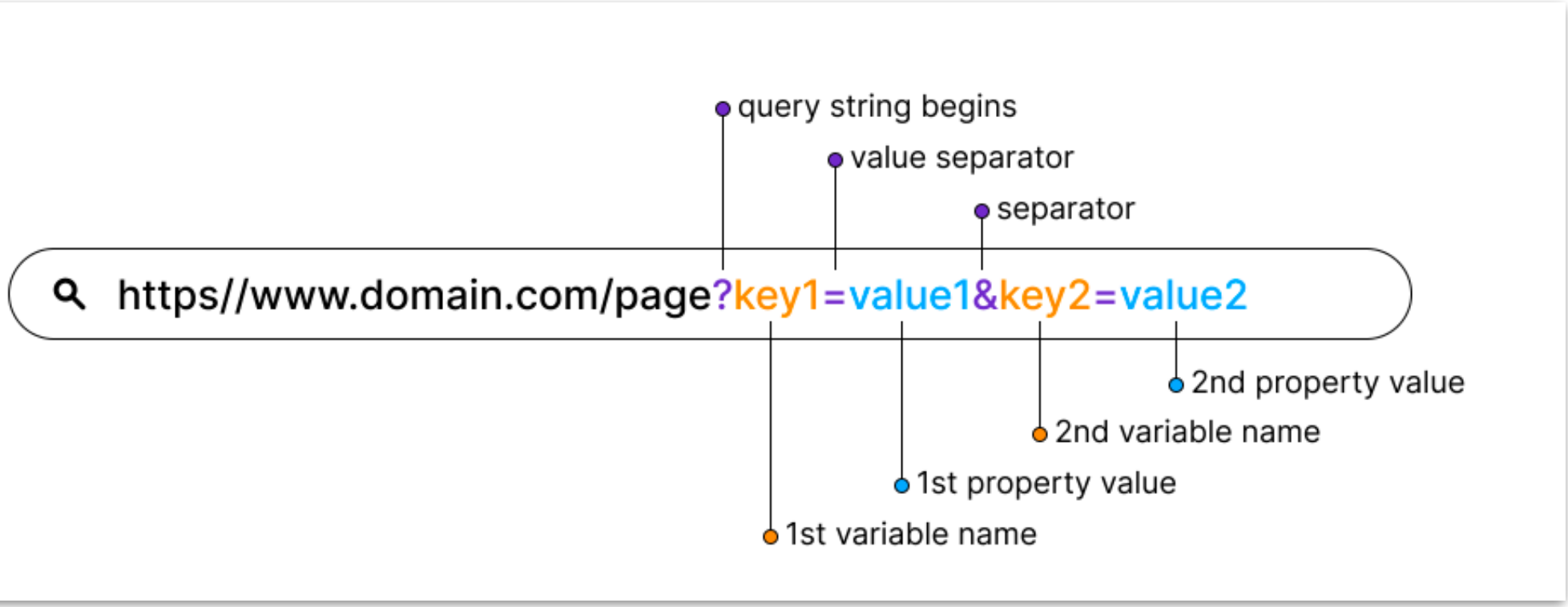
A tiny, 1x1 pixel image that is loaded when you visit a website or open an email. It's often camouflaged with the background colour or made transparent, so it is impossible to see.

Can be used to send information & kickstart code execution

Used to get information such as:

- OS
- Device (mobile, laptop, desktop)
- Type of browser
- Screen resolution
- Time it was opened
- Your actions during the session
- IP address



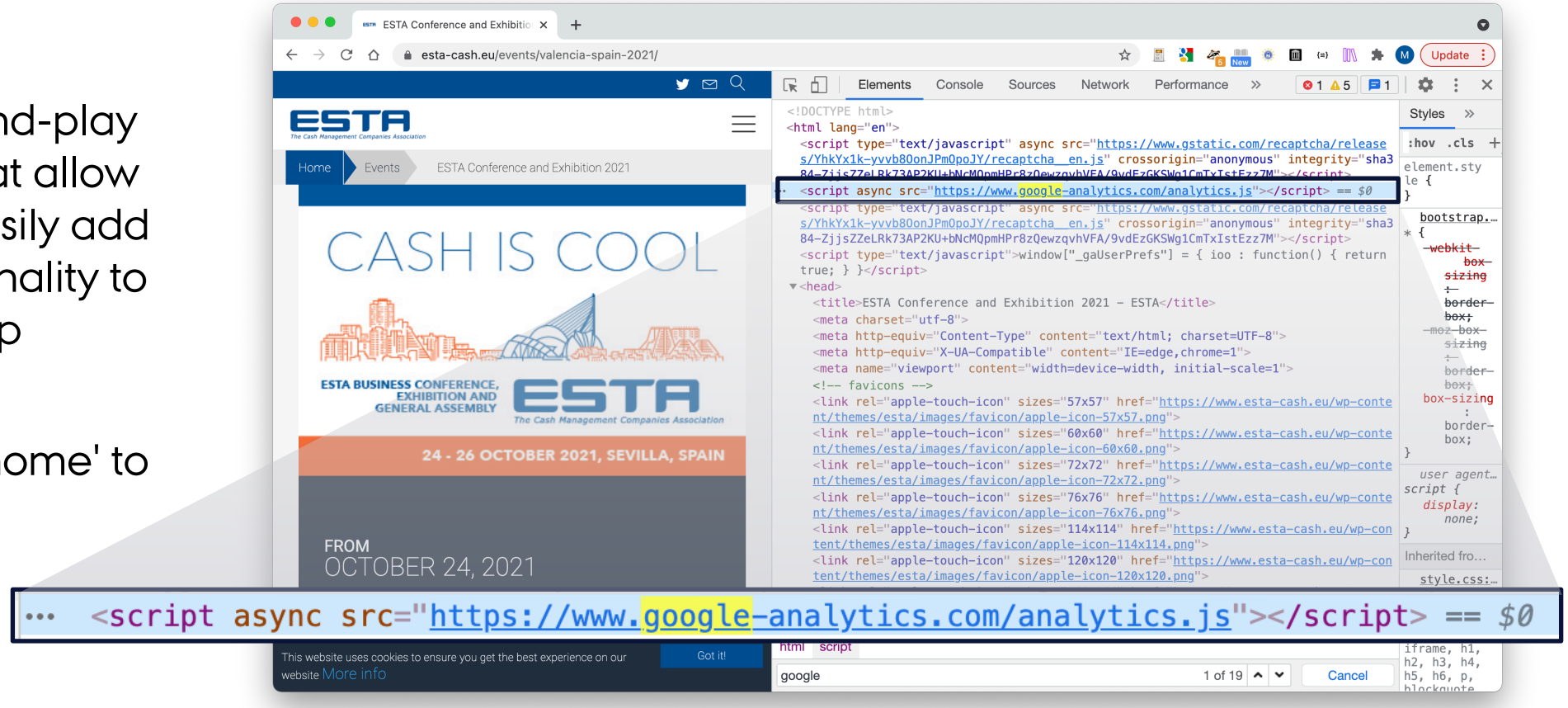



```
                <img
src="https://certify.alexametrics.com/atrk.gif?
account=x5HVh1aUXR00Gi" style="display:none"
height="1" width="1" alt="" />
                <!-- End Alexa Certify Javascript -->
              " == $0
```

SOFTWARE DEVELOPMENT KITS

SDKs are plug-and-play code libraries that allow developers to easily add common functionality to their website/app

They often 'call home' to their developers

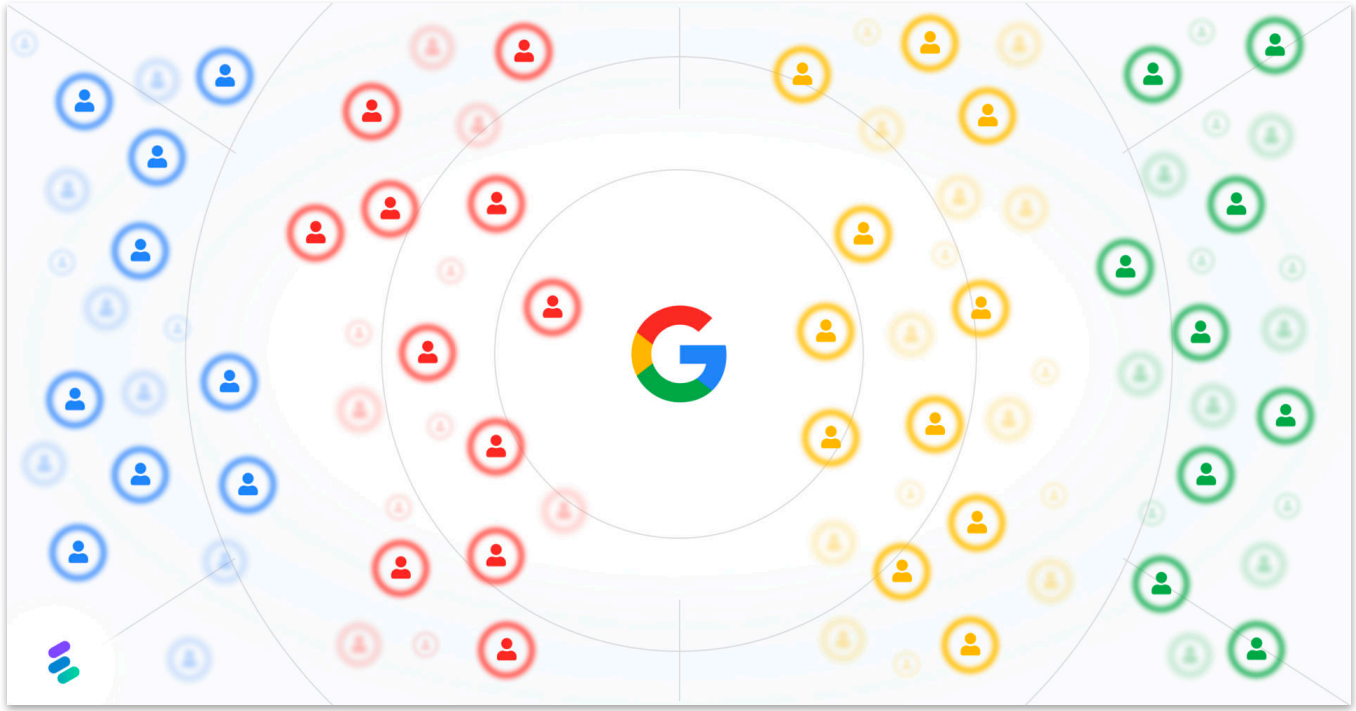


FEDERATED LEARNING OF COHORTS (TBD)

Instead of many services accessing tracking data to segment users,

Google will segment you 'locally' and signal that information to advertisers

Less 'raw' data in the world, more centralisation



Federated Learning of Cohorts (FLoC) is a new way for advertisers and sites to show relevant ads without tracking individuals across the web. FLoC was developed with the idea to protect individuals' privacy by placing them in a large crowd — a cohort — of thousands of people with similar recent browsing activity without any of them being individually identified.

https://privacysandbox.com/proposals/floc

REGULATION: EPD & GDPR

ePrivacy Directive (2002/2009)

> 'Member States shall ensure that the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user is only allowed on condition that the subscriber or user concerned has given his or her consent, having been provided with clear and comprehensive information, in accordance with Directive 95/46/EC,

General Data Protection Regulation

> (11) 'consent' of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her;

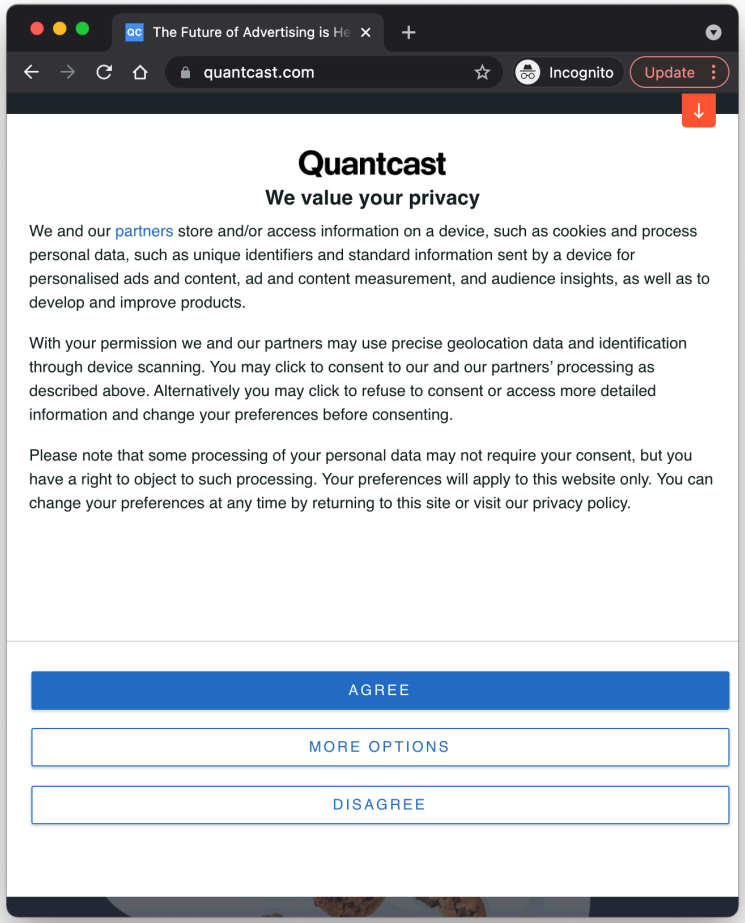
Freely given: You are not unduly influenced in your decision

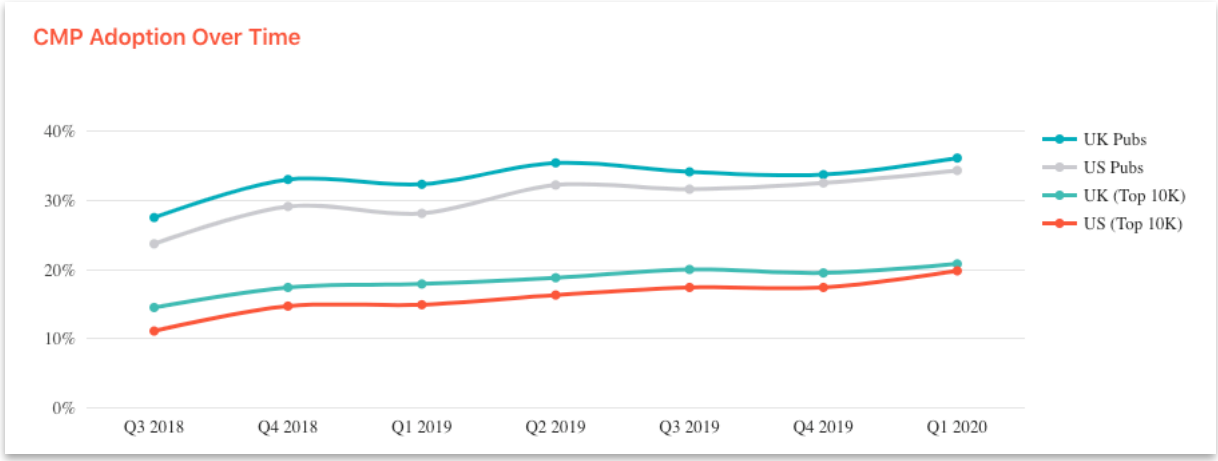
Specific: You have granular choices

Informed: You have enough information to know what you are deciding about

Unambiguous: There is no question that you have given your answer

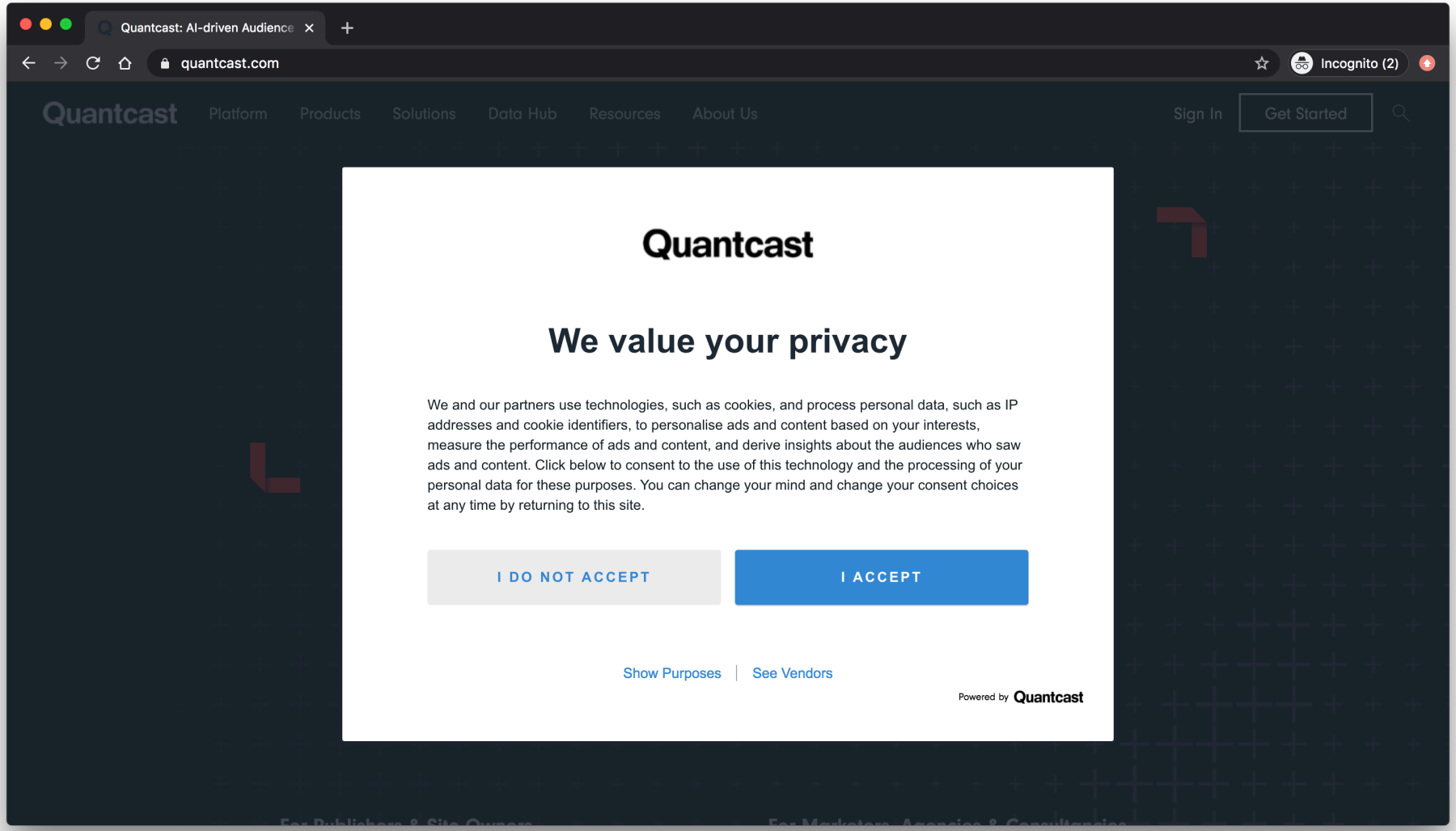
CONSENT MANAGEMENT PLATFORMS



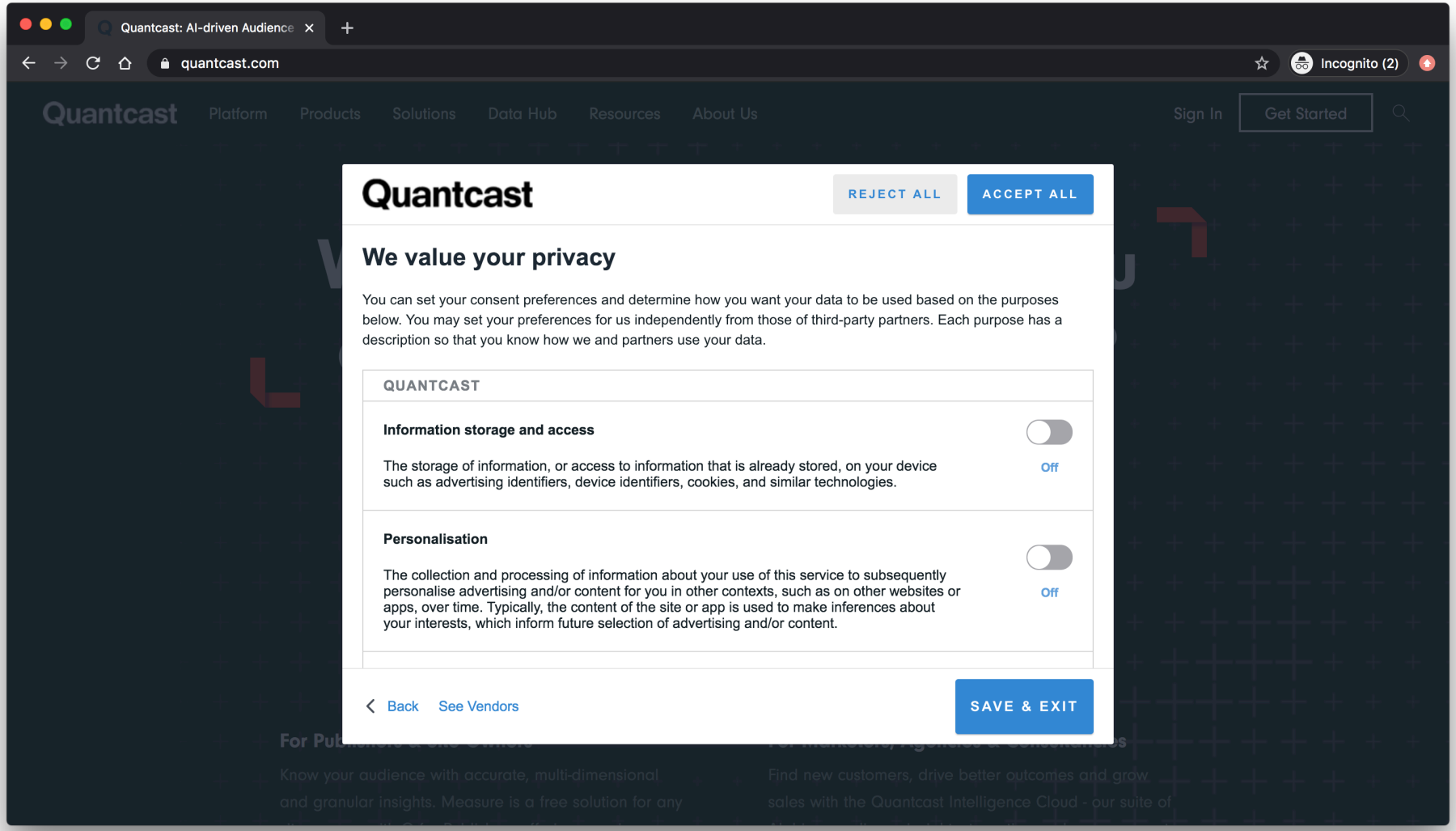


https://www.kevel.co/cmp/

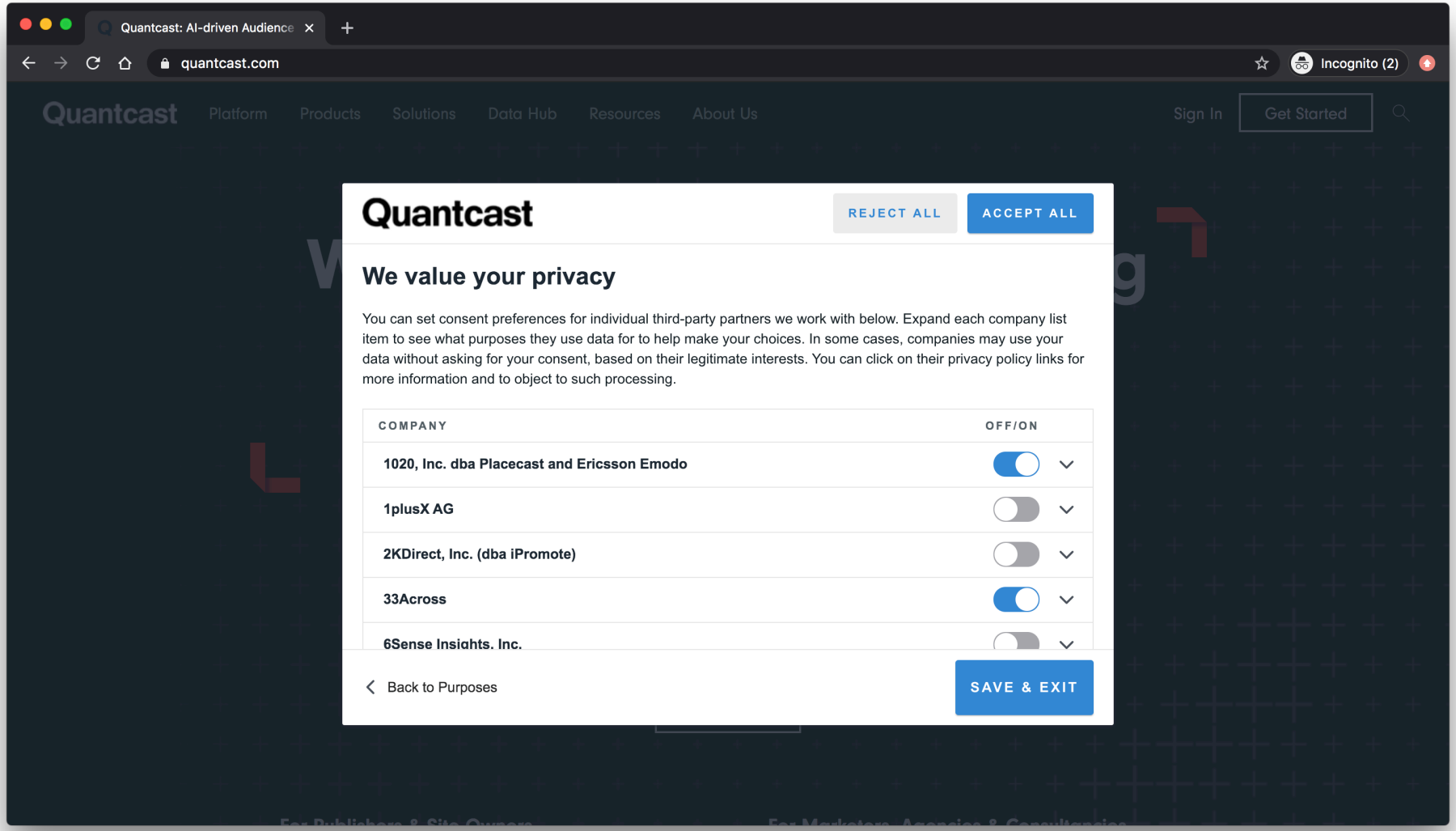
BULK


Quantcast: AI-driven Audience
quantcast.com
Incognito (2)
Quantcast
Platform
Products
Solutions
Data Hub
Resources
About Us
Sign In
Get Started
Quantcast
We value your privacy
We and our partners use technologies, such as cookies, and process personal data, such as IP addresses and cookie identifiers, to personalise ads and content based on your interests, measure the performance of ads and content, and derive insights about the audiences who saw ads and content. Click below to consent to the use of this technology and the processing of your personal data for these purposes. You can change your mind and change your consent choices at any time by returning to this site.
I DO NOT ACCEPT
I ACCEPT
Show Purposes
See Vendors
Powered by Quantcast

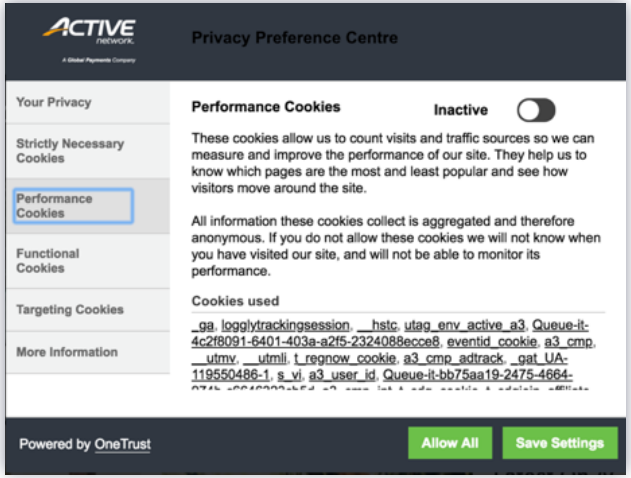
PURPOSE


Quantcast: AI-driven Audience
quantcast.com
Incognito (2)
Quantcast
Platform
Products
Solutions
Data Hub
Resources
About Us
Sign In
Get Started
Quantcast
REJECT ALL
ACCEPT ALL
We value your privacy
You can set your consent preferences and determine how you want your data to be used based on the purposes below. You may set your preferences for us independently from those of third-party partners. Each purpose has a description so that you know how we and partners use your data.
QUANTCAST
Information storage and access
The storage of information, or access to information that is already stored, on your device such as advertising identifiers, device identifiers, cookies, and similar technologies.
Off
Personalisation
The collection and processing of information about your use of this service to subsequently personalise advertising and/or content for you in other contexts, such as on other websites or apps, over time. Typically, the content of the site or app is used to make inferences about your interests, which inform future selection of advertising and/or content.
Off
Back
See Vendors
SAVE & EXIT

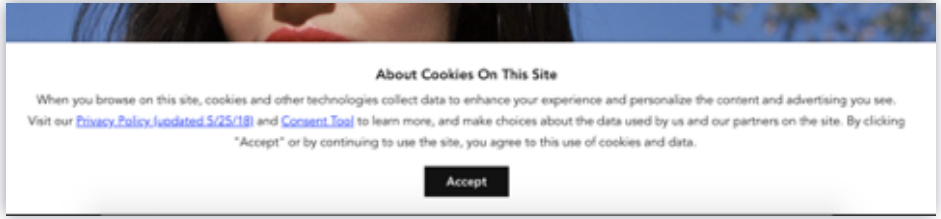
VENDORS


Quantcast: AI-driven Audience
quantcast.com
Incognito (2)
Quantcast
Platform
Products
Solutions
Data Hub
Resources
About Us
Sign In
Get Started
Quantcast
REJECT ALL
ACCEPT ALL
We value your privacy
You can set consent preferences for individual third-party partners we work with below. Expand each company list item to see what purposes they use data for to help make your choices. In some cases, companies may use your data without asking for your consent, based on their legitimate interests. You can click on their privacy policy links for more information and to object to such processing.
COMPANY
OFF/ON
1020, Inc. dba Placecast and Ericsson Emodo
1plusX AG
2KDirect, Inc. (dba iPromote)
33Across
6Sense Insights, Inc.
Back to Purposes
SAVE & EXIT

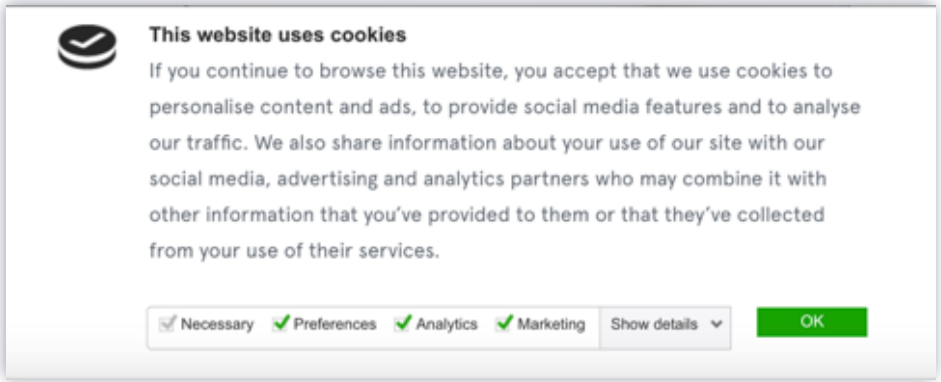
OneTrust



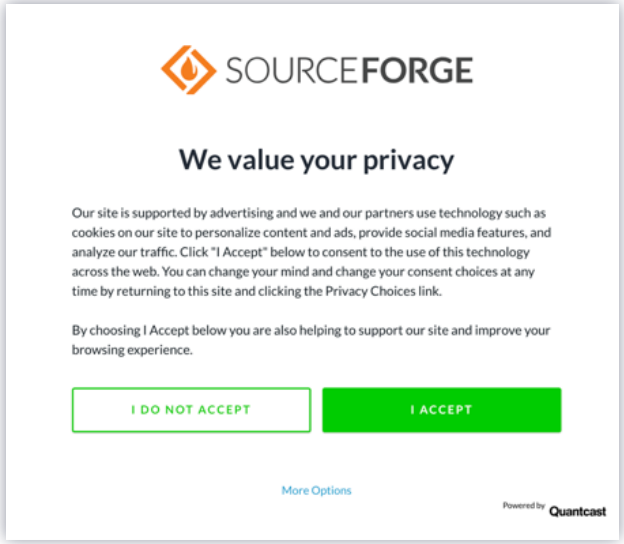
CrownPeak



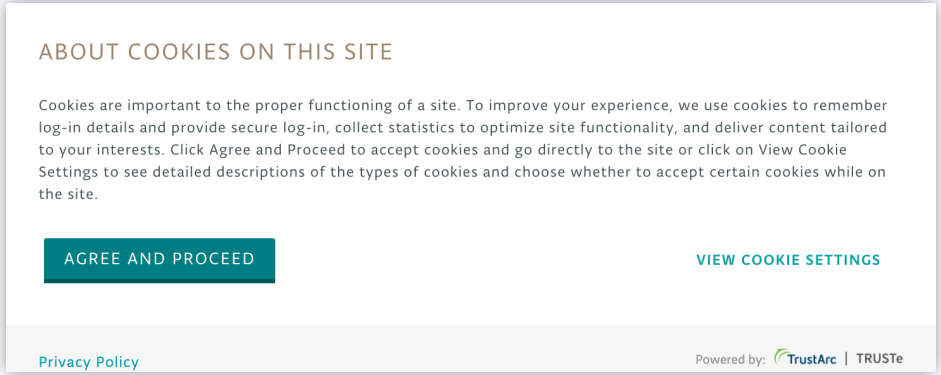
CookieBot



Quantcast



TrustArc



LARGE SCALE NON-COMPLIANCE

1. No optional boxes preticked
2. Reject all as easy as Accept all
3. Consent is explicit

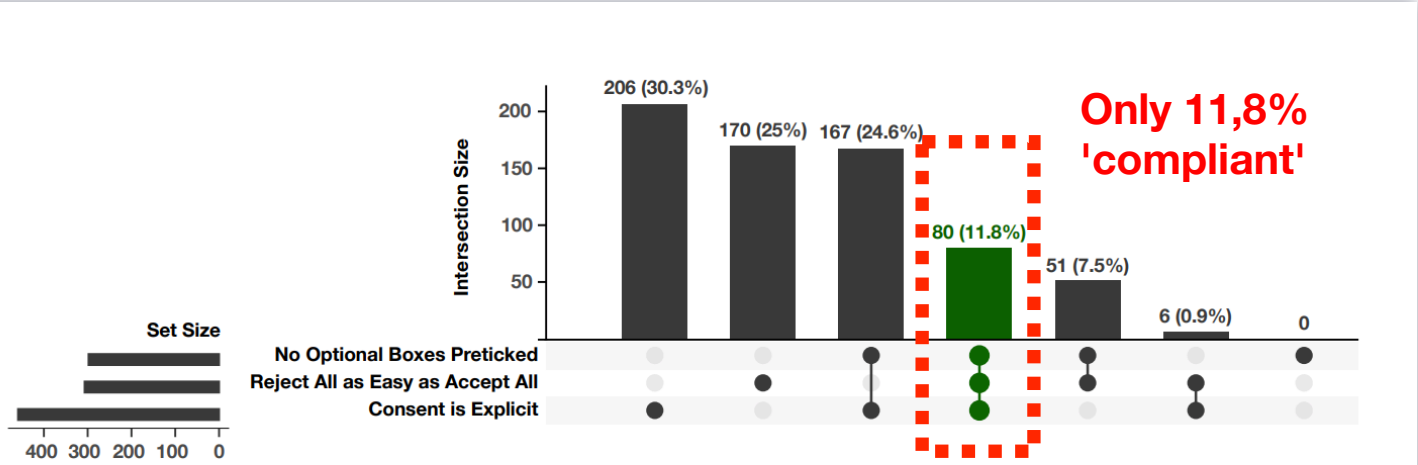


Figure 2. UpSet diagram [16, 36] of sites by adherence to three core conditions of EU law. Sites meeting all three in green.

CMP	Sites	Median vendors (low./upp. quartiles)	Explicit/implicit consent	Banner/barrier	Preticked options	Minimum compliance
Cookiebot	12.5% (85)	104 (61, 232)	45/40	78/7	64 (75.3%)	2 (5.6%)
Crownpeak	12.2% (83)	38.5 (18.8, 132.3)	46/37	52/31	67 (80.7%)	0 (0%)
OneTrust	24.3% (165)	58 (26.5, 104.5)	47/118	158/7	108 (65.4%)	3 (1.8%)
QuantCast	41% (279)	542 (542, 542)	279/0	132/147	90 (32.3%)	73 (26.2%)
TrustArc	10% (68)	87 (38, 152)	42/26	26/42	53 (77.9%)	2 (2.9%)
all	**680**	**315 (58, 542)**	**459/221**	**446/234**	**382 (56.2%)**	**80 (11.8%)**

Table 1. Key statistics on scraped CMPs.

MAIN TAKEAWAYS

- Tracking is more than just cookies, but relies on a broad suite of technologies
- Implementing these technologies is easy, but difficult to understand
- Individual countermeasures are laborious and ineffective
- Regulation already exists; widespread non-compliance in surveillance industry
- Enforcement and competition law are the key battlegrounds