



ePrivo.eu

An online privacy observatory

Pere Barlet & Ismael Castell

Universitat Politècnica de Catalunya (UPC)

ESNOG 30 – 26 Oct. 2023

Third-party trackers

GOBIERNO DE COALICION >

El PSOE y Sumar alcanzan un acuerdo de Gobierno que incluye la semana laboral de 37,5 horas

El pacto para la legislatura contempla también los impuestos a la banca y las grandes energéticas, el aumento del parque público de vivienda y la ampliación de permisos retribuidos por nacimiento



Pedro Sánchez y Yolanda Díaz, durante la firma del acuerdo, este martes.
Foto: SAMUEL SÁNCHEZ | Video: EPV

PAULA CHOUZA CARLOS E. CUÉ JOSÉ MARCOS

Madrid · 24 oct 2023 · 08:43 | Actualizado: 24 OCT 2023 · 21:50 CEST

274

La coalición da un paso decisivo para la investidura de Pedro Sánchez, a la espera de cerrar un pacto con los demás aliados de la mayoría. El PSOE y Sumar han sellado su acuerdo de

POLITICA

La intrahistoria del pacto PSOE-Sumar: "matices de matices" hasta la madrugada y sin concreciones para otros pactos "en cascada"

PREMIUM

RAÚL PIÑA

@_raulpina_
Madrid

Actualizado Miércoles,
25 octubre 2023 ·
09:26



Sánchez y Díaz sellan un acuerdo sin mención a la amnistía ni a Cataluña



Pedro Sánchez y Yolanda Díaz, en la presentación del pacto de Gobierno PSOE-Sumar. ALBERTO DI LOLLI

Coalición PSOE y Sumar
el pacto de Gobierno

Suele decirse que lo que no se nombra no existe. Quizás sea eso lo que explique que el pacto de Gobierno entre el PSOE y

Third-party trackers

For what purposes my information is used and who uses it?

This Site uses its own and other entities cookies, in order to access and use your information for the below purposes. If you do not agree with any of these purposes, you may customize your choices below.

We and the companies that collaborate with us will use your information obtained through cookies. To know the collaborating companies that incorporate their cookies on our website, such as advertisers, advertising operators and intermediaries, you can access through the button **See our partners**. You can set your consent preferences separately for each of the mentioned partners.

Additional information: You can know the complete information about the use of cookies, their configuration, origin, purposes and rights in our [Cookies Policy](#).

You allow the use of cookies for the following purposes:

+ Storage and access to information	Disagree	Agree
+ Select basic ads	Disagree	Agree
+ Create a personalised ads profile	Disagree	Agree
+ Select personalised ads	Disagree	Agree
+ Create a personalised content profile	Disagree	Agree
+ Select personalised content	Disagree	Agree
+ Measure ad performance	Disagree	Agree
+ Measure content performance	Disagree	Agree
+ Apply market research to generate audience insights	Disagree	Agree
+ Develop and improve products	Disagree	Agree
+ Sharing data and profiles for analysis and personalised advertising from advertisers for our advertising campaigns	Disagree	Agree
+ Actively scan device characteristics for identification	Disagree	Agree
+ Sharing data and profiles for analysis and personalised advertising from advertisers and advertising companies on the Internet.	Disagree	Agree
+ Use precise geolocation data	Disagree	Agree

By giving consent to the purposes above, you also allow this website and its partners to operate the following data processing: Ensure security, prevent fraud, and debug, Link different devices, Match and combine offline data sources, Receive and use automatically-sent device characteristics for identification, and Technically deliver ads or content

[View our partners](#)

PRIVACY MANAGEMENT BY DIDOMI

Disagree to all Agree to all

Select partners for Prisa

You can set your consent preferences for every partner listed below individually. Click on a partner name to get more information on what it does, what data it is collecting and how it is using it.

All partners	Block	Authorize
+ 152 Media LLC IAB TCF	Block	Authorize
+ 1Agency IAB TCF	Block	Authorize
+ 1plusX AG IAB TCF	Block	Authorize
+ 2KDirect, Inc. (dba iPromote) IAB TCF	Block	Authorize
+ 33Across IAB TCF	Block	Authorize
+ 3Q GmbH IAB TCF	Block	Authorize
+ 42 Ads GmbH IAB TCF	Block	Authorize
+ 6Sense Insights, Inc. IAB TCF	Block	Authorize
+ 7Hops.com Inc. (ZergNet) IAB TCF	Block	Authorize

[View user information](#)

PRIVACY MANAGEMENT BY DIDOMI

Save

Select partners for EIMundo

You can set your consent preferences for every partner listed below individually. Click on a partner name to get more information on what it does, what data it is collecting and how it is using it.

All partners	Block	Authorize
#		
+ 152 Media LLC IAB TCF	Block	Authorize
+ 1Agency IAB TCF	Block	Authorize
+ 1plusX AG IAB TCF	Block	Authorize
+ 2KDirect, Inc. (dba iPromote) IAB TCF	Block	Authorize
+ 33Across IAB TCF	Block	Authorize
+ 3Q GmbH IAB TCF	Block	Authorize
+ 42 Ads GmbH IAB TCF	Block	Authorize

[View user information](#)

PRIVACY MANAGEMENT BY DIDOMI

Save

Third-party trackers

For what purposes my information is used and who uses it?

This Site uses its own and other entities cookies, in order to access and use your information for the below purposes. If you do not agree with any of these purposes, you may customize your choices below.

We and the companies that collaborate with us will use your information obtained through cookies. To know the collaborating companies that incorporate their cookies on our website, such as advertisers, advertising operators and intermediaries, you can access through the button **See our partners**. You can set your consent preferences separately for each of the mentioned partners.

Additional information: You can know the complete information about the use of cookies, their configuration, origin, purposes and rights in our [Cookies Policy](#).

You allow the use of cookies for the following purposes:

+ Storage and access to information	Disagree	Agree
+ Select basic ads	Disagree	Agree
+ Create a personalised ads profile	Disagree	Agree
+ Select personalised ads	Disagree	Agree
+ Create a personalised content profile	Disagree	Agree
+ Select personalised content	Disagree	Agree
+ Measure ad performance	Disagree	Agree
+ Measure content performance	Disagree	Agree
+ Apply market research to generate audience insights	Disagree	Agree
+ Develop and improve products	Disagree	Agree
+ Sharing data and profiles for analysis and personalised advertising from advertisers for our advertising campaigns	Disagree	Agree
+ Actively scan device characteristics for identification	Disagree	Agree
+ Sharing data and profiles for analysis and personalised advertising for advertisers and advertising companies on the Internet.	Disagree	Agree
+ Use precise geolocation data	Disagree	Agree

By giving consent to the purposes above, you also allow this website and its partners to operate the following data processing: Ensure security, prevent fraud, and debug, Link different devices, Match and combine offline data sources, Receive and use automatically-sent device characteristics for identification, and Technically deliver ads or content

[View our partners](#)

PRIVACY MANAGEMENT BY DIDOMI

Disagree to all Agree to all

Select partners for Prisa

You can set your consent preferences for every partner listed below individually. Click on a partner name to get more information on what it does, what data it is collecting and how it is using it.

All partners	Block	Authorize
+ 152 Media LLC IAB TCF	Block	Authorize
+ 1Agency IAB TCF	Block	Authorize
+ 1plusX AG IAB TCF	Block	Authorize
+ 2KDirect, Inc. (dba iPromote) IAB TCF	Block	Authorize
+ 33Across IAB TCF	Block	Authorize
+ 3Q GmbH IAB TCF	Block	Authorize
+ 42 Ads GmbH IAB TCF	Block	Authorize
+ 6Sense Insights, Inc. IAB TCF	Block	Authorize
+ 7Hops.com Inc. (ZergNet) IAB TCF	Block	Authorize

+ View user information

PRIVACY MANAGEMENT BY DIDOMI

Save

1033 “partners”



Select partners for EIMundo

You can set your consent preferences for every partner listed below individually. Click on a partner name to get more information on what it does, what data it is collecting and how it is using it.

All partners	Block	Authorize
#		
+ 152 Media LLC IAB TCF	Block	Authorize
+ 1Agency IAB TCF	Block	Authorize
+ 1plusX AG IAB TCF	Block	Authorize
+ 2KDirect, Inc. (dba iPromote) IAB TCF	Block	Authorize
+ 33ACROSS IAB TCF	Block	Authorize
+ 3Q GmbH IAB TCF	Block	Authorize
+ 42 Ads GmbH IAB TCF	Block	Authorize

+ View user information

PRIVACY MANAGEMENT BY DIDOMI

Save

1128 “partners”

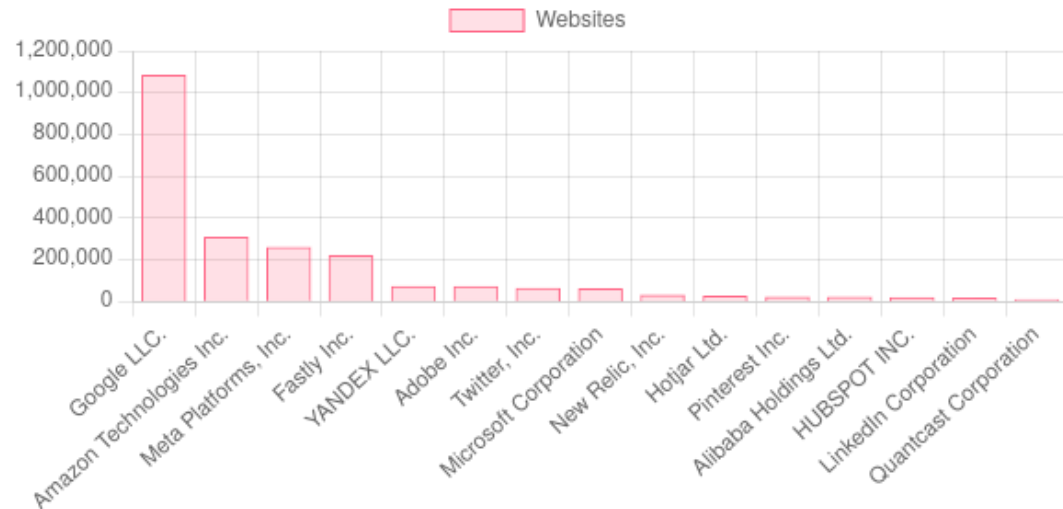
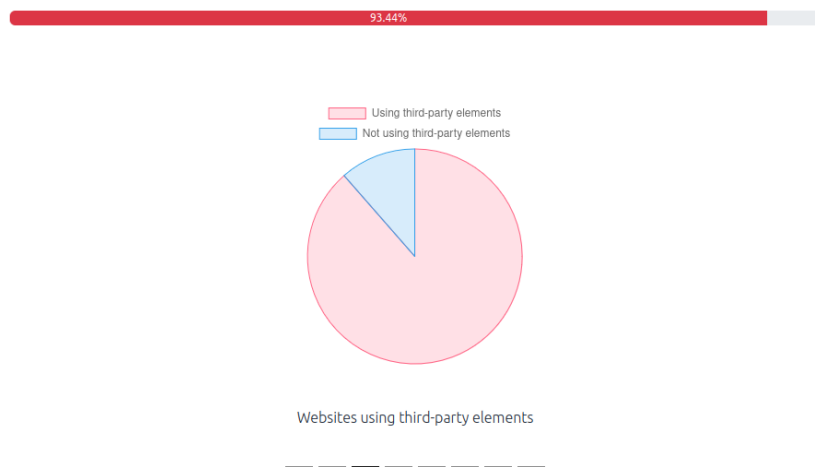
“Partners” (starting with A: 150+)

+A Million AdsIAB TCF	+AdElement Media Solutions Pvt LtdIAB TCF	+Admixer EU GmbHIAB TCF	+adsquare GmbHIAB TCF	+ADYOUlike SAIAB TCF	+Anzu Virtual Reality LTDIAB TCF	+Audience Solutions S.A.IAB TCF
+A.MobiIAB TCF		+Admo.tv (Clickon)IAB TCF	+ADSTOURS SASIAB TCF	+Adzymic Pte LtdIAB TCF	+Apester LtdIAB TCF	+Audience Trading Platform Ltd.IAB TCF
+AA INTERNET-MEDIA LtdIAB TCF	+Aderize, Inc.IAB TCF	+Adnami ApsIAB TCF	+AdsWizz Inc.IAB TCF	+Agate Systems LimitedIAB TCF	+Appier PTE LtdIAB TCF	+AudienceProject ApsIAB TCF
+Aarki, Inc.IAB TCF	+Adevinta Spain S.L.U.IAB TCF	+adnanny.com SLUIAB TCF	+Adsyield InclIAB TCF	+agof - daily campaign factsIAB TCF	+Appstock LTD.IAB TCF	+AudienceProject Measurement AmziAB TCF
+AAX LLCIAB TCF	+Adex (Virtual Minds GmbH)IAB TCF	+ADNOW I RENODO MEDIA LTDIAB TCF	+Adtarget Teknoloji A.S.IAB TCF	+agof studiesIAB TCF	+appTV Ltd.IAB TCF	+Audiencerate LTDIAB TCF
+Accorp Sp. z o.o.IAB TCF	+ADFI MEDIA SLIAB TCF	+Adnuntius ASIAB TCF	+ADTARGET.ME UABIAB TCF	+AiCreo S.r.l.IAB TCF	+Appush TechnologiesIAB TCF	+AudienceRun corpiAB TCF
+AccountInsight LtdIAB TCF	+Adform A/SIAB TCF	+Adobe Advertising CloudIAB TCF	+Adtelligent Inc.IAB TCF	+AIDEM Technologies LTDIAB TCF	+Aptivio InclIAB TCF	+Audiens S.r.l.IAB TCF
+Active Agent (Virtual Minds GmbH)IAB TCF	+AdGear Technologies, Inc.IAB TCF	+Adobe Audience Manager, Adobe Experience PlatformIAB TCF	+AdTheorent, InclIAB TCF	+Airnow Media LtdIAB TCF	+AragoIAB TCF	+Audiennz AGIAB TCF
+ACTV8, Inc.IAB TCF	+AdheseIAB TCF	+Adobe Audience Manager, Adobe Experience Platform	+AdTiming Technology Company LimitedIAB TCF	+Akamai	+Arbeitsgemeinschaft Media-AnalyseIAB TCF	+AuDigentIAB TCF
+AcuityAds Inc.IAB TCF	+adhood.comIAB TCF	+AdomikIAB TCF	+Adtriba GmbHIAB TCF	+AlgoriX Technology Pte. Ltd.,IAB TCF	+Arcanor Bilgi Teknolojileri ve Hizmetleri A.Ş.IAB TCF	+audio content & control GmbHIAB TCF
+AcxiomIAB TCF	+AdikteevIAB TCF	+Adpone SLIAB TCF	+ADTTTRIBUTION InclIAB TCF	+AlkimilIAB TCF	+ArcSpan Technologies, Inc.IAB TCF	+Audiohook CorporationIAB TCF
+Ad Alliance GmbHIAB TCF	+AdInMo LTDIAB TCF	+adQueryIAB TCF	+ADUXIAB TCF	+Allegro sp. z o.o.IAB TCF	+Arcspire LimitedIAB TCF	+AUDIOMOB LTDIAB TCF
+ad6medialIAB TCF	+ADITION (Virtual Minds GmbH)IAB TCF	+AdQuiver Media SLIAB TCF	+advanced store GmbHIAB TCF	+Alliance Gravity Data MediaIAB TCF	+AREIA (ADLOOP)IAB TCF	+AudionIAB TCF
+Adacado Technologies Inc. (DBA Adacado)IAB TCF	+Adjust Digital A/SIAB TCF	+ADRENALIAIAB TCF	+ADventori SASIAB TCF	+Alphonso Inc. IAB TCF	+ArkeeroIAB TCF	+Automatic AdsIAB TCF
+adality GmbHIAB TCF	+Adjust GmbHIAB TCF	+Adrino Mobile Sp. z o.o.IAB TCF	+Adventure MediaIAB TCF	+Always-Computing SASIAB TCF	+ARKHEUSIAB TCF	+Avantis Video LtdIAB TCF
+Adasta Media S.r.l.IAB TCF	+Adkernel LLCIAB TCF	+adrule mobile GmbHIAB TCF	+AdverlineIAB TCF	+Amazon Ad ServerIAB TCF	+ARMIS SASIAB TCF	+Avocet Systems LimitedIAB TCF
+adbouncer Werbeagentur GmbHIAB TCF	+Adlane LTDIAB TCF	+Ads Interactive Ltd.IAB TCF	+Adverticum cPlc.IAB TCF	+Amazon AdvertisingIAB TCF	+Arpeely Ltd.IAB TCF	+AWIN AGIAB TCF
+adbility media GmbHIAB TCF	+Adloox SAIAB TCF	+Adserve.zone / Artworx ASIAB TCF	+Advertising.Tech FZ-LLCIAB TCF	+Amnet GmbHIAB TCF	+Arrivalist Co.IAB TCF	+Axel Springer Teaser Ad GmbHIAB TCF
+AdClear GmbHIAB TCF	+Adludio Ltd.IAB TCF	+AdServing Factory srlIAB TCF	+Advertly AB (publ)IAB TCF	+Amobee Inc.IAB TCF	+ArtChaos s.r.o.IAB TCF	+Axiom Media ConnectIAB TCF
+AdColony, Inc.IAB TCF	+ADMAN - Phaistos Networks, S.A.IAB TCF	+AdViewIAB TCF	+Advertly AB (publ)IAB TCF	+Amplified IntelligenceTechnologiesIAB TCF	+Artefact Deutschland GmbHIAB TCF	+AXIS CORP LTDIAB TCF
+AddAppttr GmbHIAB TCF	+ADman Interactive SLUIAB TCF	+Advisible ABIAB TCF	+Advertly AB (publ)IAB TCF	+ANINPRO-CREATIVE, S.L.IAB TCF	+Aryel s.r.l.IAB TCF	+Axonix LTDIAB TCF
+AdDefend GmbHIAB TCF	+adMarketplace, Inc.IAB TCF	+Adwatch	+Advertly AB (publ)IAB TCF	+Aniview LTDIAB TCF	+Ask LocalIAB TCF	+Azerion Holding B.V.IAB TC
+AdDefend GmbHIAB TCF	+ADMAXIAB TCF	+Adsolutions BVIAB TCF	+ADWAYS SASIAB TCF	+AnonymisedIAB TCF	+AttributyIAB TCF	+ >850 more partners!!
+AddrelevanceIAB TCF	+AdMaxim LimitedIAB TCF	+AdsonicalIAB TCF	+adWWMGIAB TCF	+Anteriad, LLCIAB TCF	+Audience NetworkIAB TCF	
+Adelaide Metrics InclIAB TCF	+admetrics GmbHIAB TCF	+Adspirit GmbHIAB TCF	+Adxperience SASIAB TCF	+AntVoiceIAB TCF		

Third-party trackers

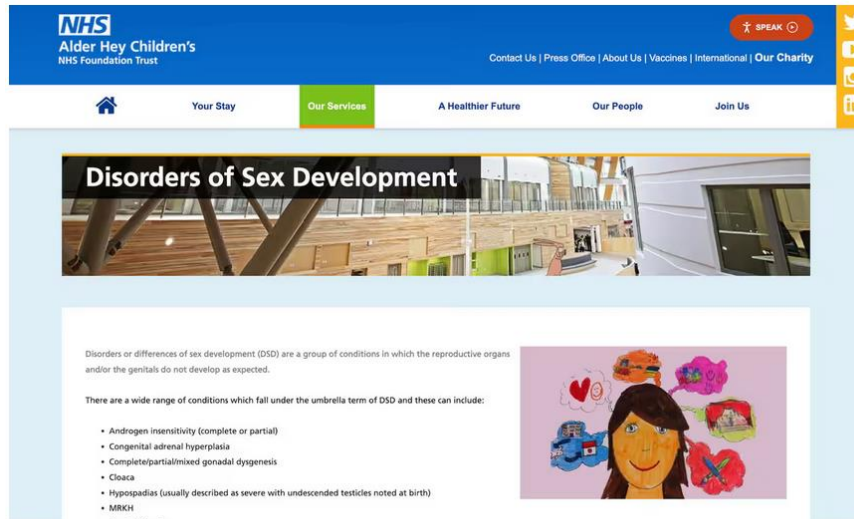
- Third-party trackers are present in >90% of the websites
- They can easily recover our *browsing history* (at least)
- They can track us even if we delete cookies, use private mode, DNT, VPN, etc.

Percentage of domains using web tracking



Third-party trackers

NHS data breach: trusts shared patient details with Facebook without consent



📷 A page about sexual development disorders on Alder Hey Children's Hospital's website, which shared details of the browsing with Facebook via the Meta Pixel.

Observer investigation reveals Meta Pixel tool passed on private details of web browsing on medical sites

NHS trusts are sharing intimate details about patients' medical conditions, appointments and treatments with [Facebook](#) without consent and despite promising never to do so.

An *Observer* investigation has uncovered a covert tracking tool in the websites of 20 [NHS](#) trusts which has for years collected browsing information and shared it with the tech giant in a major breach of privacy.

The data includes granular details of pages viewed, buttons clicked and keywords searched. It is matched to the user's IP address - an identifier linked to an individual or household - and in many cases details of their Facebook account.

Information extracted by [Meta Pixel](#) can be used by Facebook's parent company, Meta, for its own business purposes - including improving its targeted advertising services.

Records of information sent to the firm by NHS websites reveal it includes data which - when linked to an individual - could reveal personal medical details.

It was collected from patients who visited hundreds of NHS webpages about HIV, self-harm, gender identity services, sexual health, cancer, children's treatment and more.

It also includes details of when web users clicked buttons to book an appointment, order a repeat prescription, request a referral or to complete an online counselling course. Millions of patients are potentially affected.

Data brokers

- Companies or business units whose primary source of revenue is selling user personal information in data market places [1]
- Dangers for the user
 - Profiles may be (most surely) incorrect
 - There is no easy way to discover and correct it (despite GDPR)
 - There is no easy way to know to whom our info was sold to [2,3]
 - There is no easy way to know how this is impacting us!!



[1] A. Rieke, H. Yu, D. Robinson, and J. van Hoboken, "Data Brokers in an Open Society", Open Society Foundations 2016

[2] <https://nrkbeta.no/2020/12/03/my-phone-was-spying-on-me-so-i-tracked-down-the-surveillants/>

[3] <https://techcrunch.com/2023/05/21/the-government-cant-seize-your-data-but-it-can-buy-it/>

Tracking purposes

- Purposes and implications
 - Targeted advertising
 - Price discrimination [4]
 - Insurance coverage [5]
 - Credit rating [6]
 - Phishing/identity theft (e.g. Conti Leaks)
 - Political interests (e.g. Cambridge Analytica)



Issue: September 21, 2021

IMF report suggests credit scores could soon be based on web browsing history

Dystopian future.

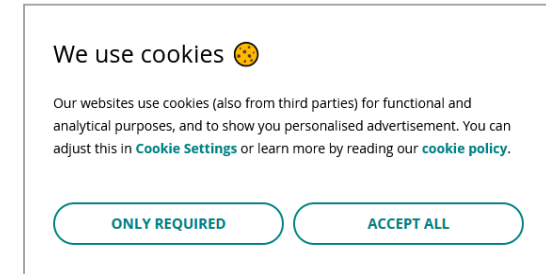
[4] J. Mikians, L. Gyarmati, V. Erramilli, and N. Laoutaris, "Crowd-assisted search for price discrimination in e-commerce: first results," in Proceedings of the ninth ACM conference on Emerging networking experiments and technologies, ser. CoNEXT '13

[5] The Economist, "Insurance data: Very personal finance," 2012. [Online]. Available: <http://www.economist.com/node/21556263>

[6] Reclaim the Net, [Online] <https://reclaimthenet.org/imf-report-suggests-credit-score-browsing-history>

Tracking methods

- Stateful methods
 - Cookies
 - HTML5 storage [7]
 - Web cache [8], DNS cache [9], Redirect cache [10]
- Stateless methods (fingerprinting)
 - Combine device properties to form a unique identifier
 - Canvas fingerprinting [11]
 - Audio fingerprinting [12]
 - 3D rendering fingerprinting [13]
 - Mouse fingerprinting [14]



	Device Fingerprinting	OS Fingerprinting	Browser Fingerprinting
IP Address	✓		
Geolocation	✓		
OS version	✓	✓	
Screen resolution	✓	✓	
Timezone	✓		
Fonts	✓	✓	
Browser version	✓	✓	✓
Browser language	✓		✓
Multitouch capabilities	✓		
Pointer capabilities	✓		
Keyboard layout	✓		
Microphone	✓		
Camera	✓		
TCP timestamp	✓		
OS architecture		✓	
System language		✓	
Color depth		✓	
Audio capabilities	✓		
CPU brand	✓		
GPU brand	✓		
System drivers	✓		
Browser instance			✓
Browser plugins			✓
Accepted media formats			✓
Cookies enabled			✓
Supercookies limitations			✓

[7] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild," in ACM CCS, 2014.
[8] G. Fleischer, "Implementing Web Tracking," Proc. Black Hat USA Conf. Briefings, pp. 1–37, Jul. 2002.
[9] A. Klein and B. Pinkas, "DNS Cache-Based User Tracking." [Online]. Available: <https://www.ndss-symposium.org/ndss-paper/dns-cache-based-user-tracking/>
[10] E. Bursztein, "Tracking users that block cookies with a http redirect," Jul. 2011.
[11] K. Mowery and H. Shacham, "Pixel Perfect : Fingerprinting Canvas in HTML 5," in Proceedings of Web 2.0 Security and Privacy (W2SP), 2012
[12] S. Englehardt and A. Narayanan, "Online Tracking: A 1-million-site Measurement and Analysis," in ACM CCS, 2016
[13] Y. Cao, S. Li, and E. Wijmans, "(Cross-)Browser Fingerprinting via OS and Hardware Level Features," NDSS Symposium, 2017
[14] M. Urpí, I. Castell and P. Barlet. "Detecting and Analyzing Mouse Tracking in the Wild," IEEE EuroS&PW, 2023

Adblockers

- Most popular protection methods
 - Block URLs using pattern lists
 - Difficult to maintain updated and block new methods
 - Easy to evade by periodically changing the hosting domains
 - Adblockers are *too* conservative
- Why are they conservative?
 - “Users want protection that does not break websites” [15]
 - “When extensions break websites, users disable them based on how important the content they are accessing is to them” [16]
- Usually do not block URLs that can impact the website functionality



[15] C. E. Wills and D. C. Uzunoglu, “What Ad Blockers Are (and Are Not) Doing,” HotWeb, 2016

[16] G. Merzdovnik, M. Huber, D. Buhov, N. Nikiforakis, S. Neuner, M. Schmiedecker, and E. Weippl, “Block Me If You Can: A Large-Scale Study of Tracker-Blocking Tools,” EuroS&P, 2017

TrackSign

- Objective: Automatically detect *unknown* web trackers
 - Known trackers under unknown URLs
 - Completely new tracking methods
- TrackSign analyzes website code instead of URLs
 - If a piece of website code is mostly present in files which we already know that perform tracking, it is most probably being used for tracking
 - We can automatically label as tracking other files using the same code!
- TrackSign explores two aspects of the code:
 - Popularity: #resources that use this piece of code
 - Dirtiness: %resources known to perform tracking
- The proposal is based on two observations:
 - Limited number of ways to obtain user information
 - New effective tracking methods will become popular at some point

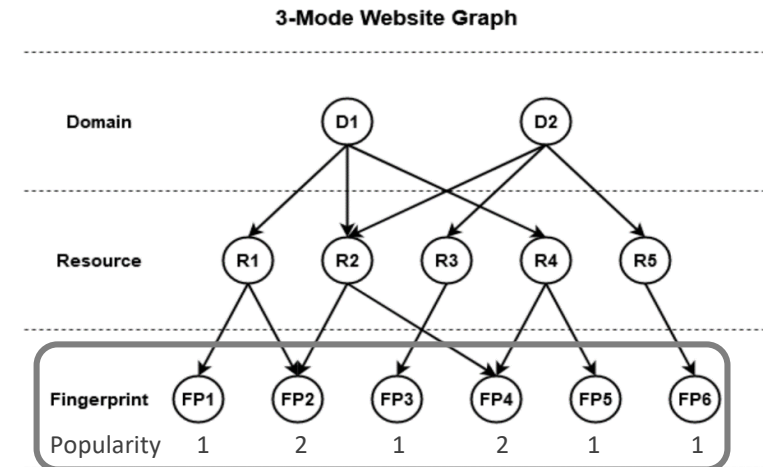
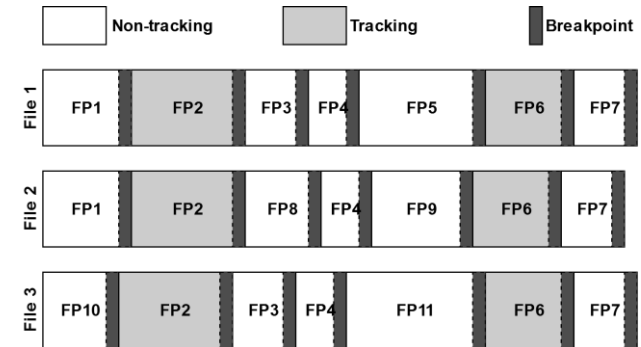
TrackSign

- Challenges

- How to partition the code?
- Same file hosted in several servers and linked by many different URLs

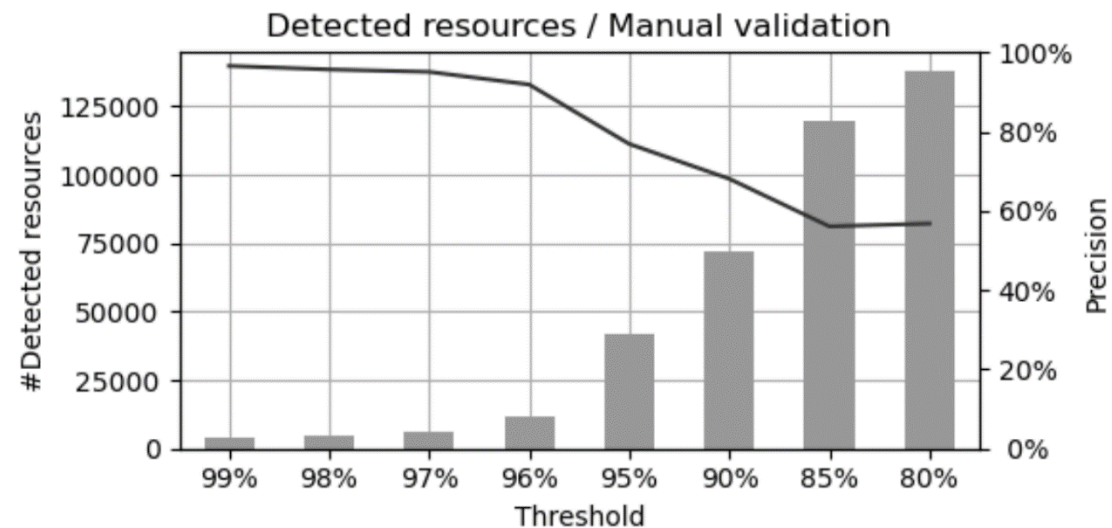
- Solution

- Rabin Fingerprinting for file partitioning
- A 3-mode network graph (tripartite)
 - Domain, Resource and Signature

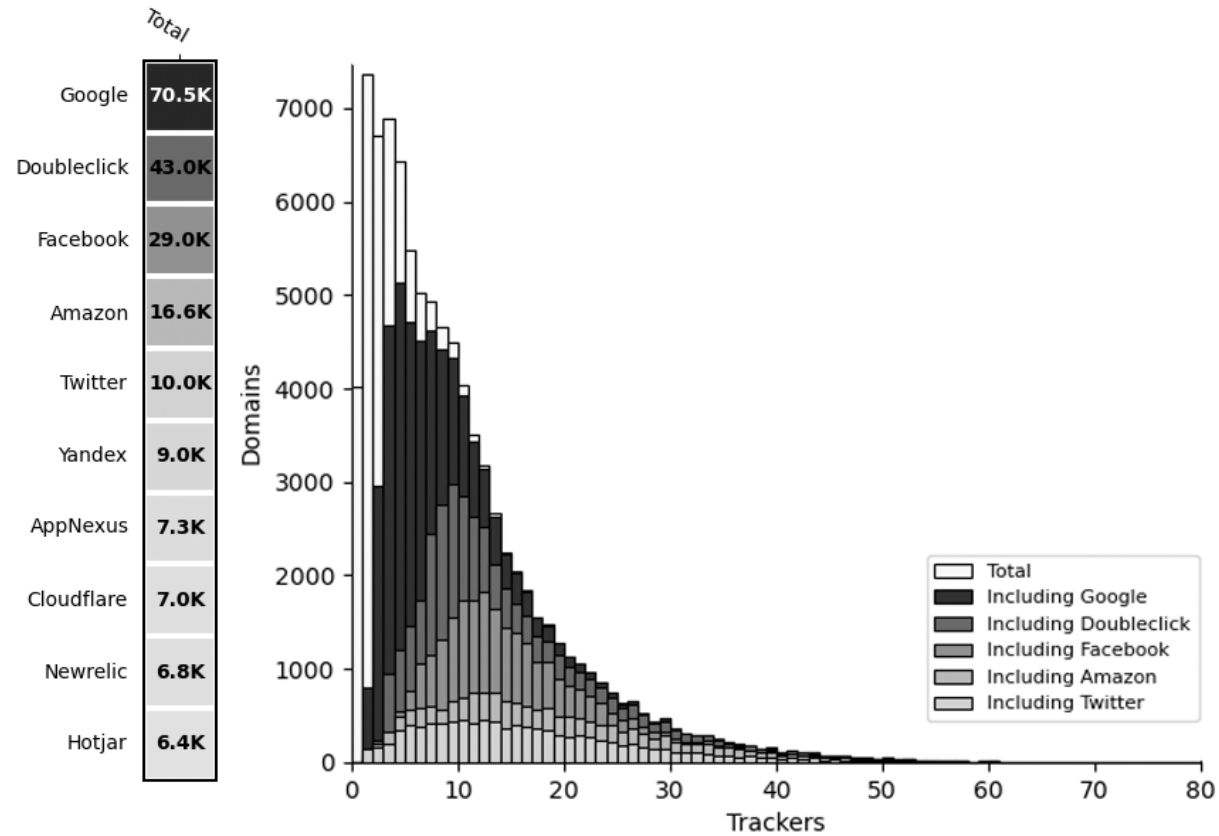


TrackSign evaluation

- Top 100.000 most popular websites
 - 5.2M of unique URLs, 933K resources, 73M code signatures
 - Manually validated about 800 files
- 96% threshold maximizes detection while maintaining accuracy (92%)
 - Found more than 12.000 new tracking resources and 300.000 new tracking URLs



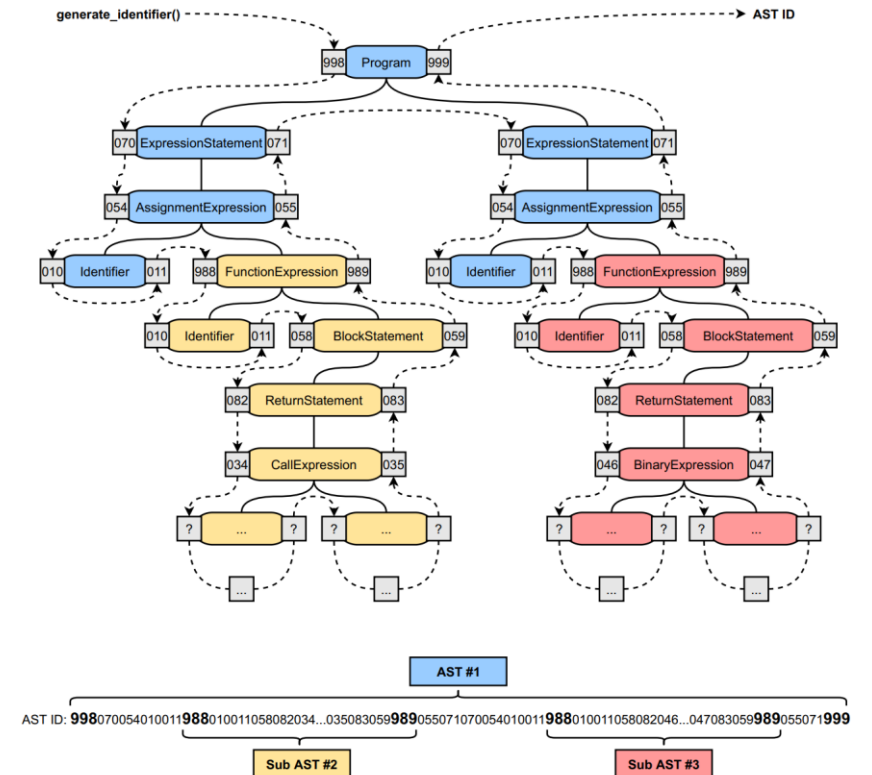
TrackSign evaluation



- 95% of websites (86.626 out of 90.637) include web tracking

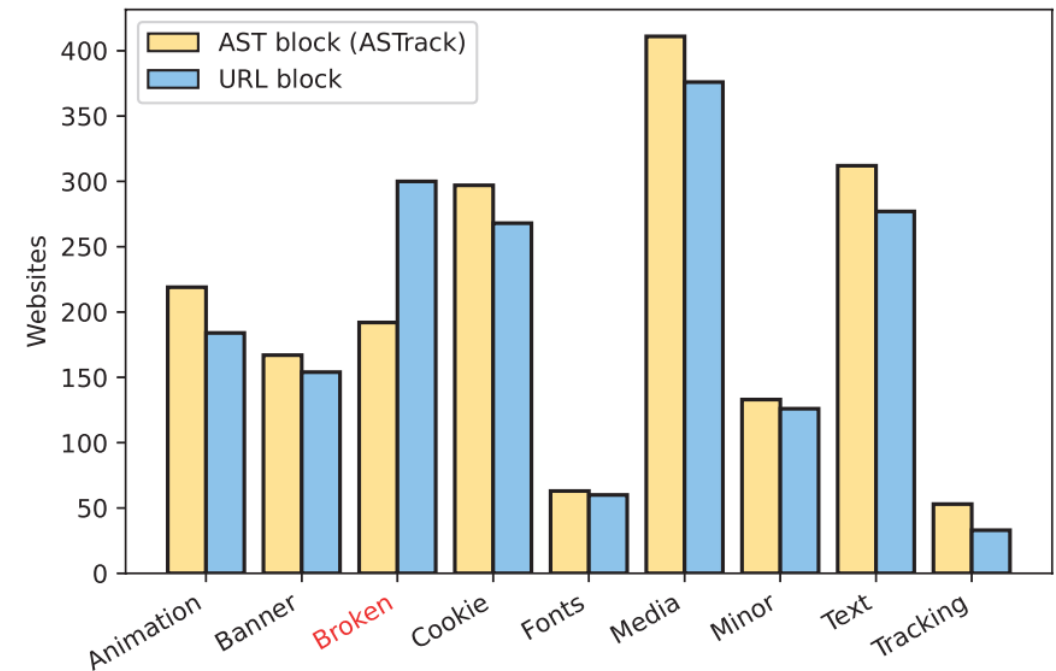
ASTrack

- Analyze code structure instead of shared code
 - We identify shared *functionality* not shared code
 - Robust against obfuscation and minimification
 - We can selectively block tracking AST better preserving functionality!
- ASTrack explores Abstract Syntax Trees (AST)
 - ASTs identified by a simple label chain



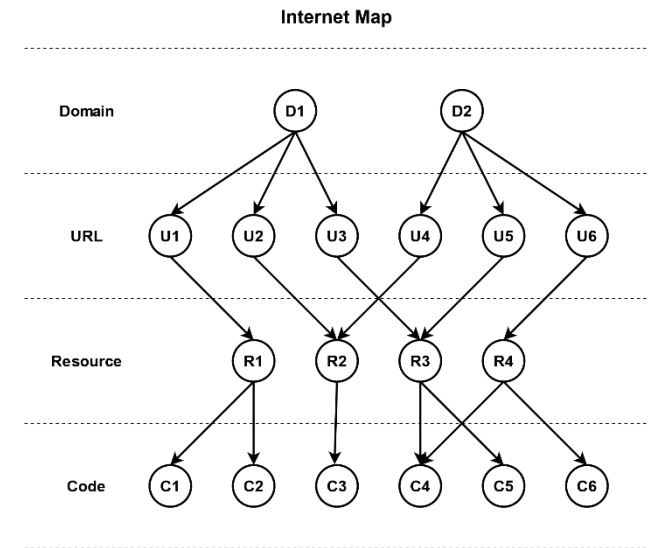
ASTrack evaluation

	Static evaluation	Dynamic evaluation
Domains		8,179
URLs		615,780
JavaScript URLs		161,593
Tracking URLs		41,274
Unique ASTs		7,015,542
Shared ASTs		2,683,586
New tracking URLs	3,409	2,183
New tracking JavaScript files	3,109	2,093
Tracking ASTs	49,453	41,114
Precision	98.52%	98.47%




- We block more than 3k URL compared to adblockers, while improving website functionality!

- New online privacy observatory
 - Data collected using TrackSign/ASTrack over 1.5 million websites
- Augmented with a new layer about URL information to obtain privacy information not included inside website code
 - Cookies
 - Certificates
- Also explores prevalence of third-party trackers



ePrivo Search History trackers Research About



Welcome to ePrivo Observatory

Search a domain to analyze with ePrivo

ePrivo is a tool designed to explore the privacy-friendliness of any given website. ePrivo explores all the internal URLs and resources, looking for web tracking systems running in the background. The detection is based in a state-of-the-art algorithm called TrackSign which is able to heuristically detect web tracking code inside JavaScript and HTML files.

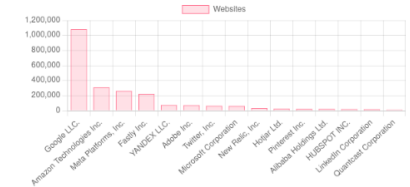
Introduce a domain...

Why is data privacy important?

Data privacy is crucial to protect individuals' personal information and prevent it from being misused. With the increasing amount of personal data being collected and shared online, it's essential to ensure that this information is kept secure and only used for its intended purposes.

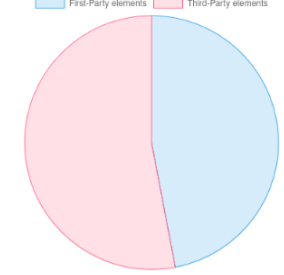
Percentage of domains using web tracking

93.44%




Company	Count
Google LLC	~1,100,000
Amazon Technologies Inc.	~300,000
Meta Platforms, Inc.	~250,000
Facebook, Inc.	~200,000
Yandex LLC	~150,000
Alibaba Inc.	~100,000
Twitter, Inc.	~50,000
Microsoft Corporation	~50,000
New York, Inc.	~50,000
Intel Ltd.	~50,000
Protonet Inc.	~50,000
Alibaba Holding Ltd.	~50,000
Walmart Inc.	~50,000
LinkedIn Corporation	~50,000
Quantum Corporation	~50,000

First-Party vs Third-Party



Legend: First-Party elements (blue), Third-Party elements (red)

Website tracking

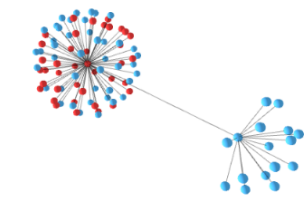


Tracking Method	Prevalence
Session cookies	25
Long living cookies	~10
Very long living cookies	~5
Javascript cookies	~5
Third party cookies	~5
Tracking cookies	~5
Font fingerprinting	~5
Canvas fingerprinting	~5
Canvas fingerprinting tag	~5
Mouse fingerprinting	~5
Webgl fingerprinting	~5
Other tracking	~5

Domain info

Domain name	elpais.com
Elements	100
Intrusion level	100%
First-Party elements	47
Third-Party elements	53

Tracking vs Third-Party vs Type



Legend: Not-Tracking (blue), Tracking (red)

Left-click: rotate, Mouse-wheel/middle-click: zoom, Right-click: pan

Conclusions

- Summary
 - ePrivo.eu can help Internet users realize the huge problem around online privacy
 - ePrivo.eu could also be useful by regulators or even web site owners
 - We can now automatically detect unknown web tracking and discover new methods
 - We can block more than adblockers by keeping website functionality in most cases
- Future work
 - Bring ePrivo.eu into a production state
 - Develop a browser extension that implements TrackSign/ASTrack
- What do we need?
 - It has proven very hard to find funding for this project...
 - We could only develop it as a side project 😞



ePrivo.eu

An online privacy observatory

Pere Barlet & Ismael Castell
{pere.barlet, ismael.castell}@upc.edu

Universitat Politècnica de Catalunya (UPC)

ESNOG 30 – 26 Oct. 2023

Backup slides

Privacy regulations

	GDPR	CCPA	PIPL
Year	2018	2020	2021
Under protection	People living in Europe	People born in California	People living in China
Target	For-profit and non-profit national or private	For-profit companies (with conditions)	For-profit and non-profit national or private
Area of effect	World	State of California	World
Restricted collection scenarios	Yes (6)	No	Yes (7)
Required consent	Yes	No	Yes
Forces service without consent	Yes	No	Yes
Opt-out available	Yes	Partially (data selling processes)	Yes
Access right	Yes	Yes	Yes
Amendment right	Yes	Yes	Yes
Deletion right	Yes	Partially	Yes
Right to be forgotten	Yes	No	Yes
Exemptions for national agencies	No	Yes	Yes
Limits data retention	Yes	No	Yes
Policies on automated algorithms	Yes	No	Yes

- Complemented with automatic labeling of web tracking methods

Fingerprinting

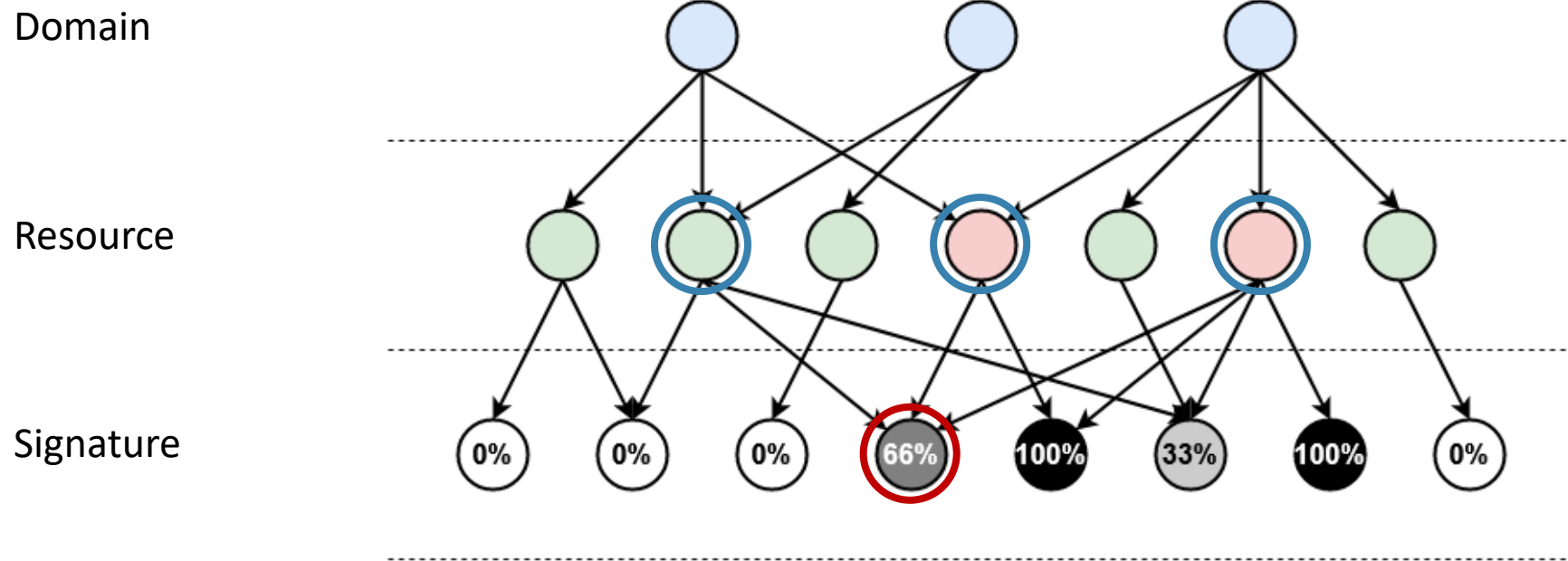
- Font fingerprinting
- Canvas fingerprinting
- Mouse fingerprinting
- WebGL fingerprinting

Cookies

- Session cookies
- Long-living cookies (less than 1 year)
- Very long-living cookies (more than 1 year)
- JavaScript cookies
- Third-party cookies
- Tracking cookies

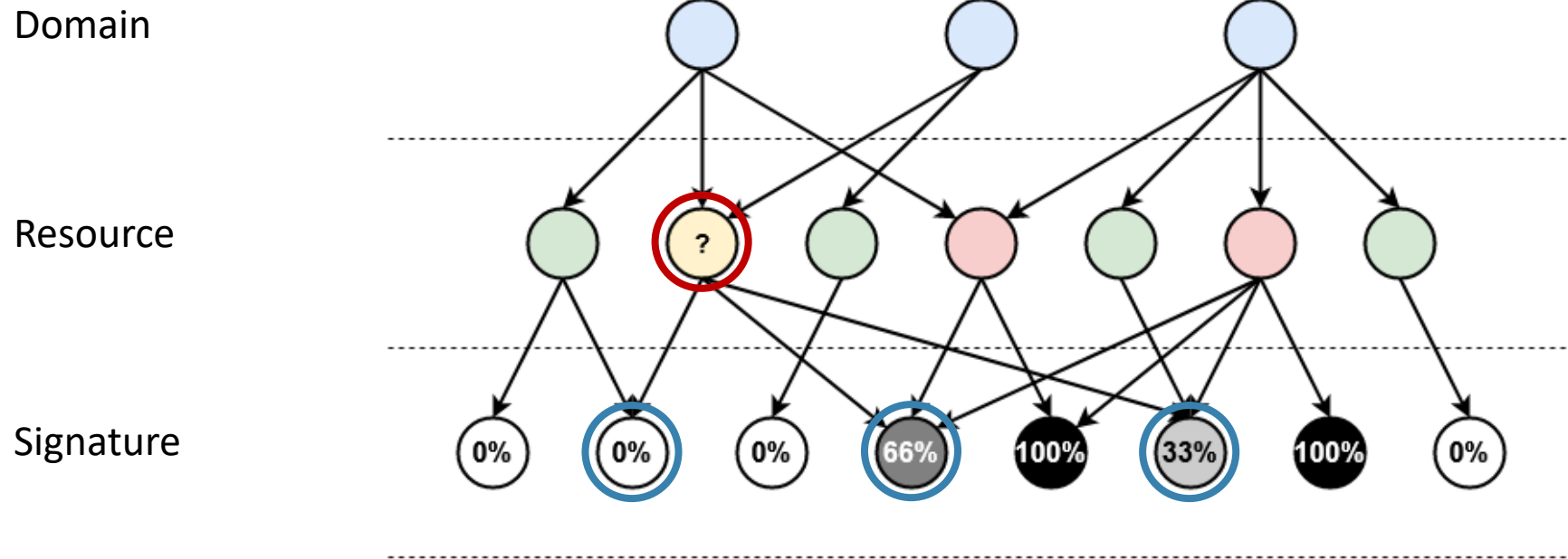
TrackSign Algorithm

Step 1: Initial labeling + fingerprint Dirt level pre-process



TrackSign Algorithm

Step 2: Backwards classification (threshold 60%)



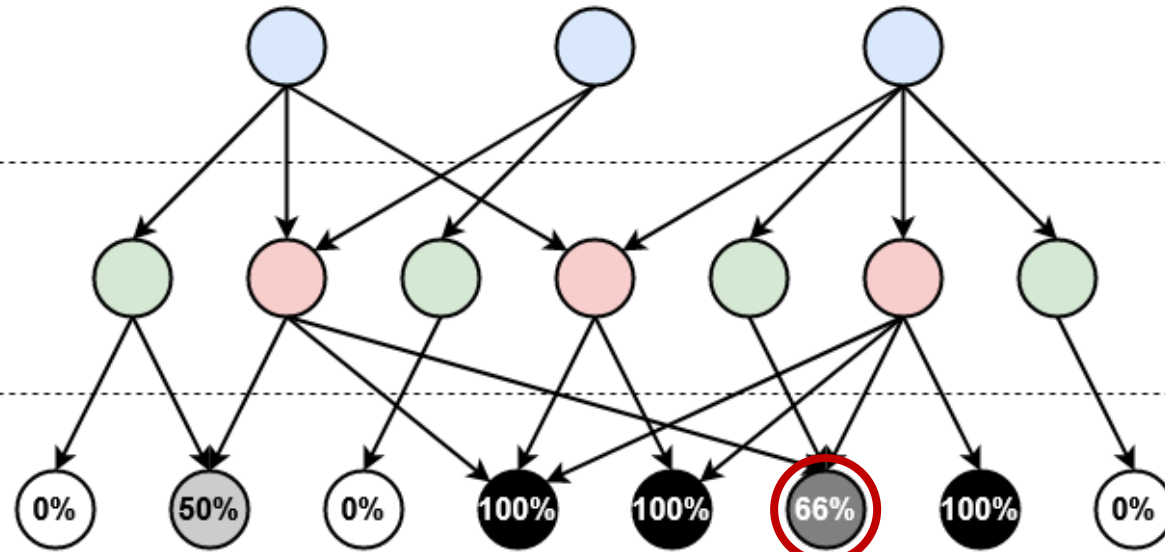
TrackSign Algorithm

Step 3: Fingerprint Dirt level update

Domain

Resource

Signature



TrackSign evaluation

- Completely new web tracking -> popularity level
- Wayback machine
- Popularity 10 -> 737 resources

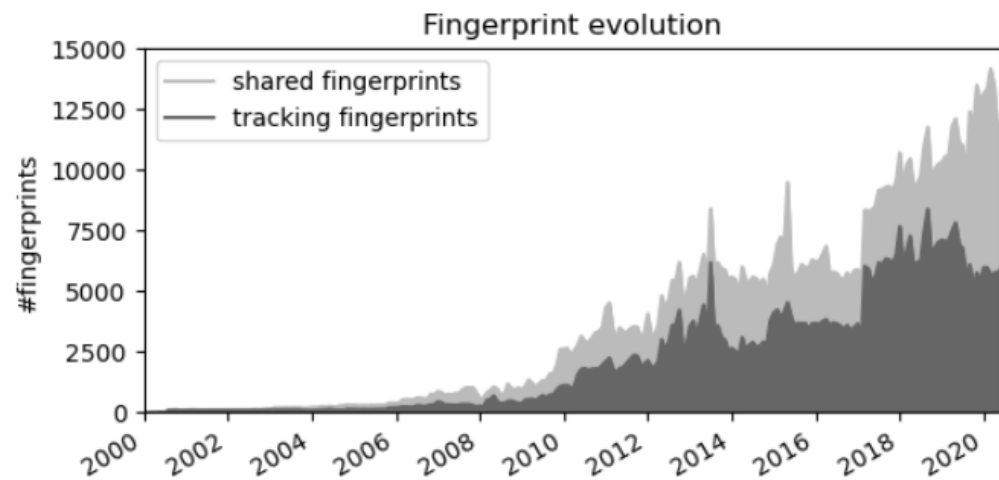
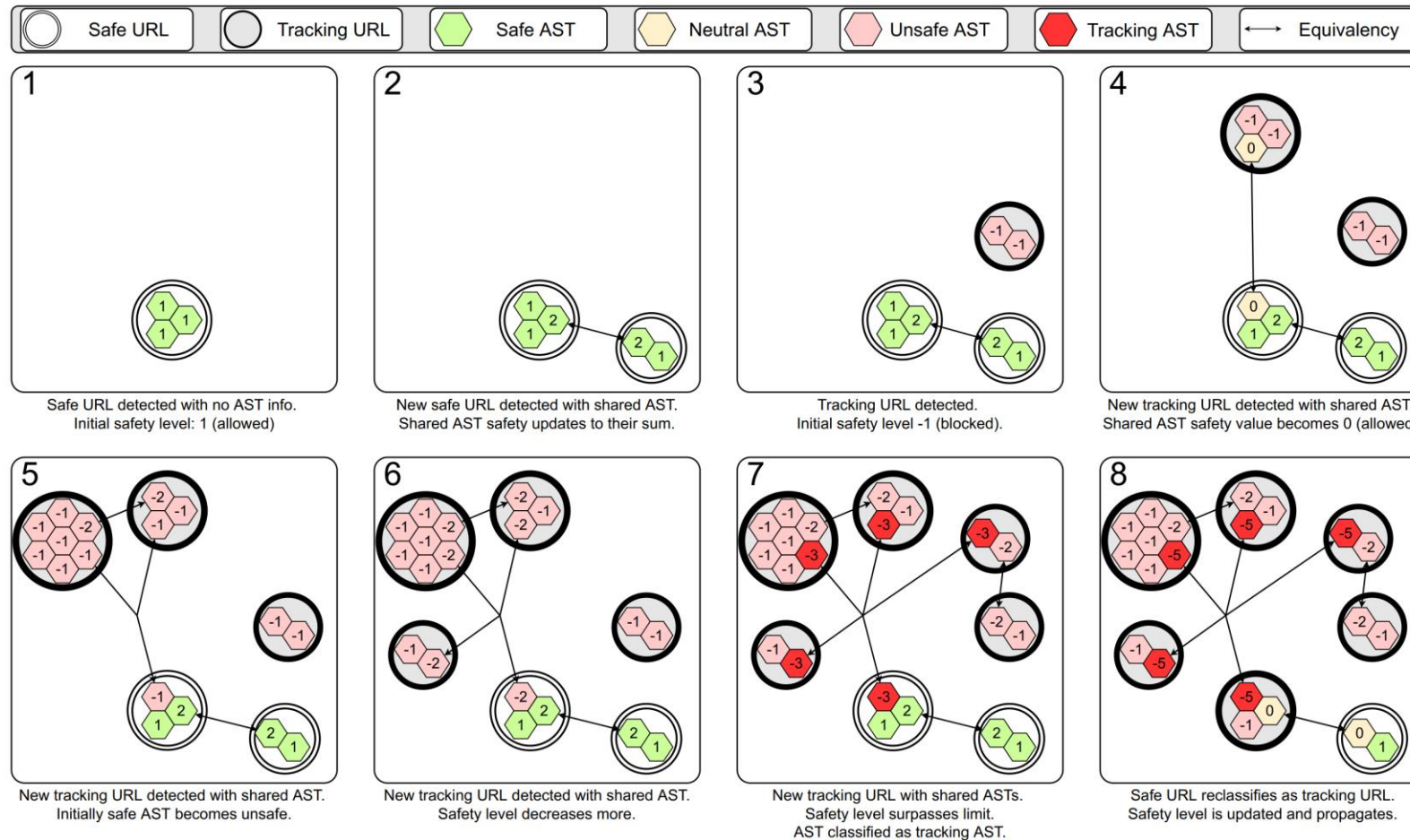


TABLE II
MOST POPULAR SERVICES

Service	Tracking	Domains	URLs	First detection
Archive.org custom libraries	No	26	127	2000-09
Archive.org analytics	Yes	54	276	2001-04
JQuery library	No	64	465	2003-02
Flash player detection library	Yes	14	14	2007-09
Google analytics (library call)	Yes	24	52	2009-06
Google tracking	Yes	29	2578	2012-04
Facebook events (library call)	Yes	10	10	2015-06
Google Tag Manager (library call)	Yes	31	45	2016-08
Google analytics	Yes	14	14	2017-01
Youtube tracking	Yes	30	403	2017-03
Zoom.us	No	21	21	2020-03

ASTrack algorithm



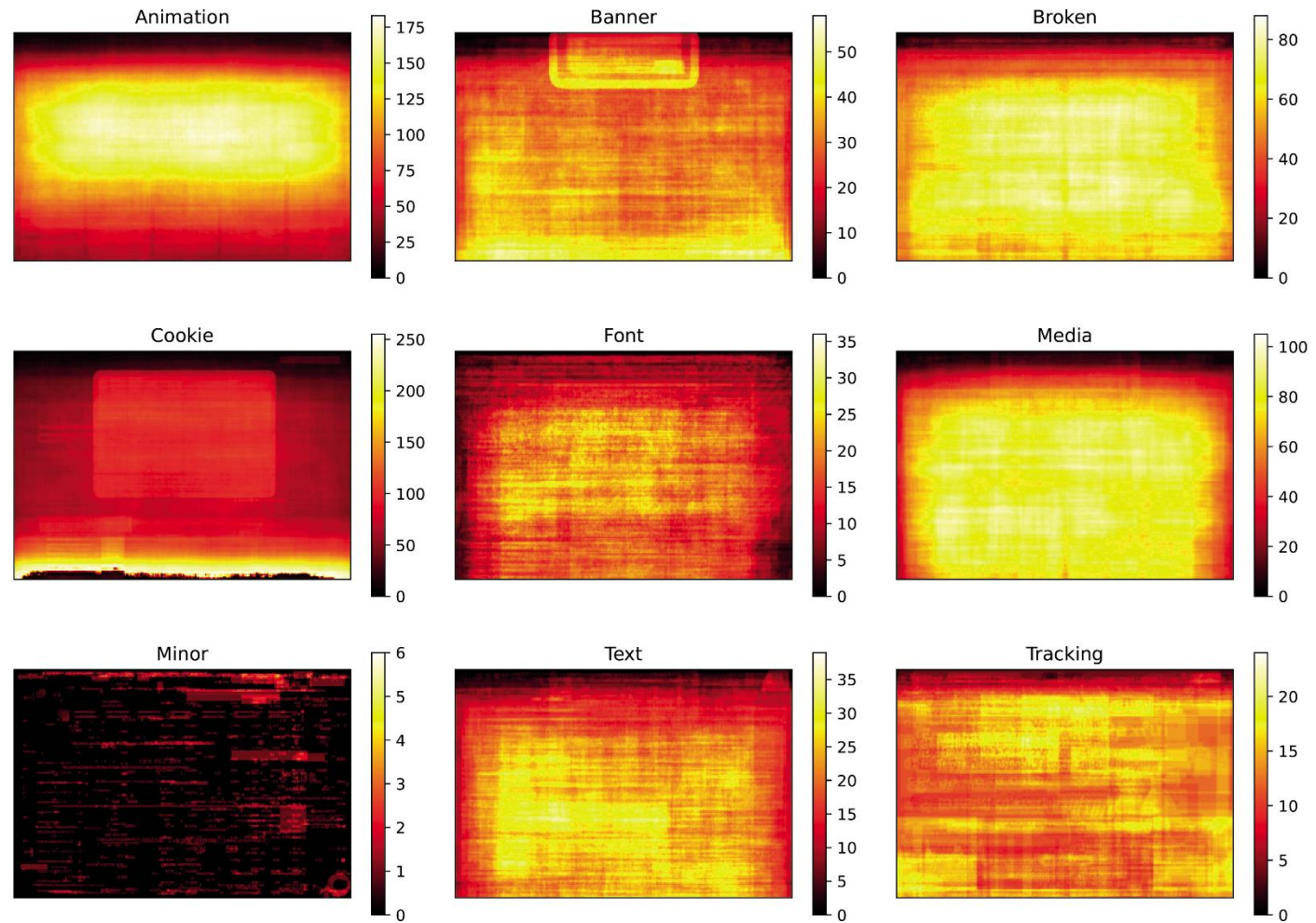
ATrack evaluation

The image shows a screenshot of the Zoom website's homepage, viewed through a browser window titled "visualizer". The website features the Zoom logo, navigation links for "SOLUTIONS", "PLANS & PRICING", "CONTACT SALES", and "RESOURCES", along with "JOIN", "HOST", "SIGN IN", and a "SIGN UP, IT'S FREE" button. The main headline reads "In this together. Keeping you connected wherever you are." Below this is a form to "Enter your work email" and a "Sign Up Free" button. A "MEETINGS" section is visible at the bottom.

Overlaid on the page is a red-bordered box containing a cookie consent message. The text inside the box reads: "Zoom uses cookies and similar technologies as strictly necessary to make our site work. We and our partners would also like to set additional cookies to enhance your use of our site, to personalize and enhance your visit to our site and to show you more relevant content and advertising. These will be set only if you accept." Below this text are three buttons: "COOKIE SETTINGS", "DECLINE COOKIES", and "ACCEPT COOKIES".

On the right side of the browser window, there is a "Filters" panel with various checkboxes: "Equal", "Different", "Text", "Media", "Animation", "Cookie", "Font", "Tracking", "Banner", "Minor", and "Broken". Below the filters is a vertical list of line numbers from 40 to 633. At the bottom of the browser window, there is a toolbar with various analysis tools: "Mask", "Vanilla", "Substitutor", "Difference", "Equal", "Different", "Text", "Media", "Animation", "Cookie", "Font", "Tracking", "Banner", "Minor", "Broken", "Save", and "Export".

ATrack evaluation



ATrack evaluation

Obfuscated code matching example #1

```
function a0_0x4172(_0x3dcddc, _0x3f9388){return a0_0x4172=  
  function(_0x5a6727, _0x4172bc){_0x5a6727=_0x5a6727-0x19d  
  ;var _0x22ac18=a0_0x5a67[_0x5a6727];return _0x22ac18;},  
  a0_0x4172(_0x3dcddc, _0x3f9388);}  
  
function u(g,O){return u=function(a,y){a=a-475;var b=m[a];  
  return b},u(g,O)}
```

Obfuscated code matching example #2

```
function(_0x210d59, _0x4d8907, _0x1f76d2, _0x4ffef){return  
  _0x210d59(_0x4d8907, _0x1f76d2, _0x4ffef);}  
  
function(api, id, key, location){return _callStorageFunction(  
  _getItem, arguments, location)}
```