

Challenges in building a private web

(without burning it all down)

Steven Englehardt

Privacy Engineer



Consumer Tech • Perspective

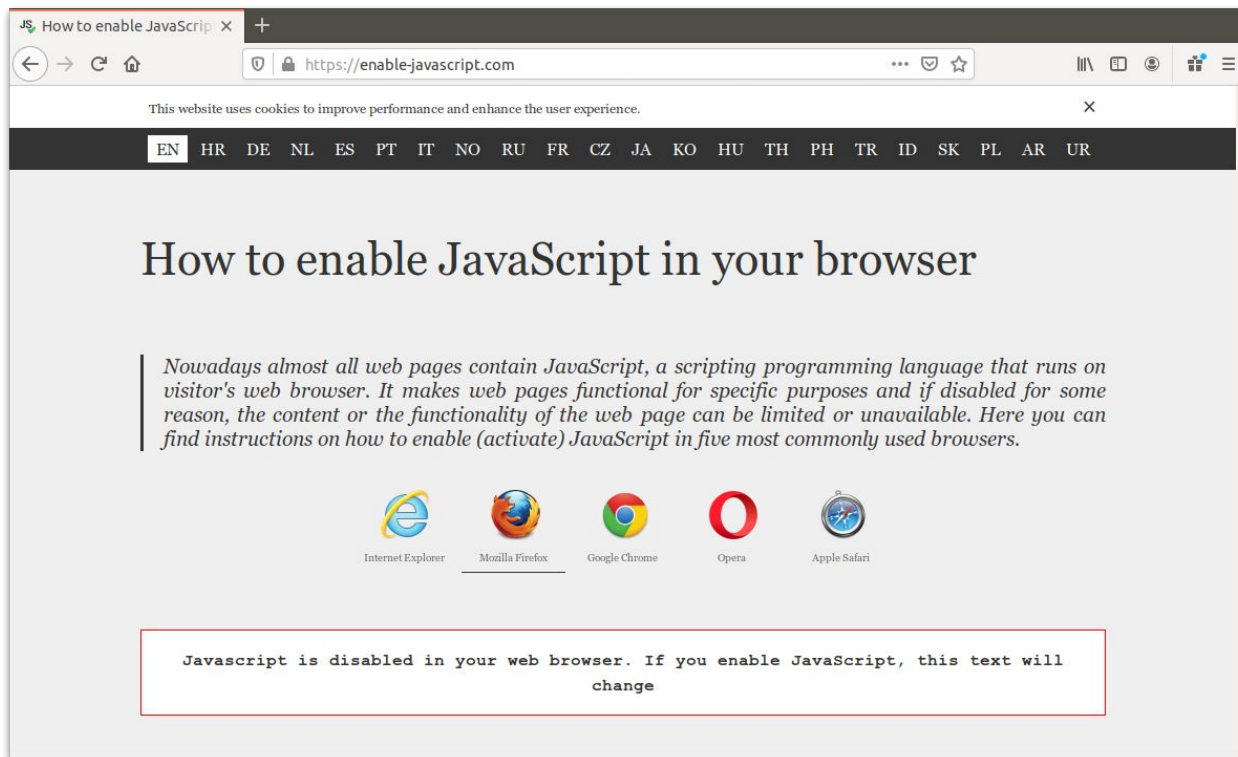
Think you're anonymous online? A third of popular websites are 'fingerprinting' you.

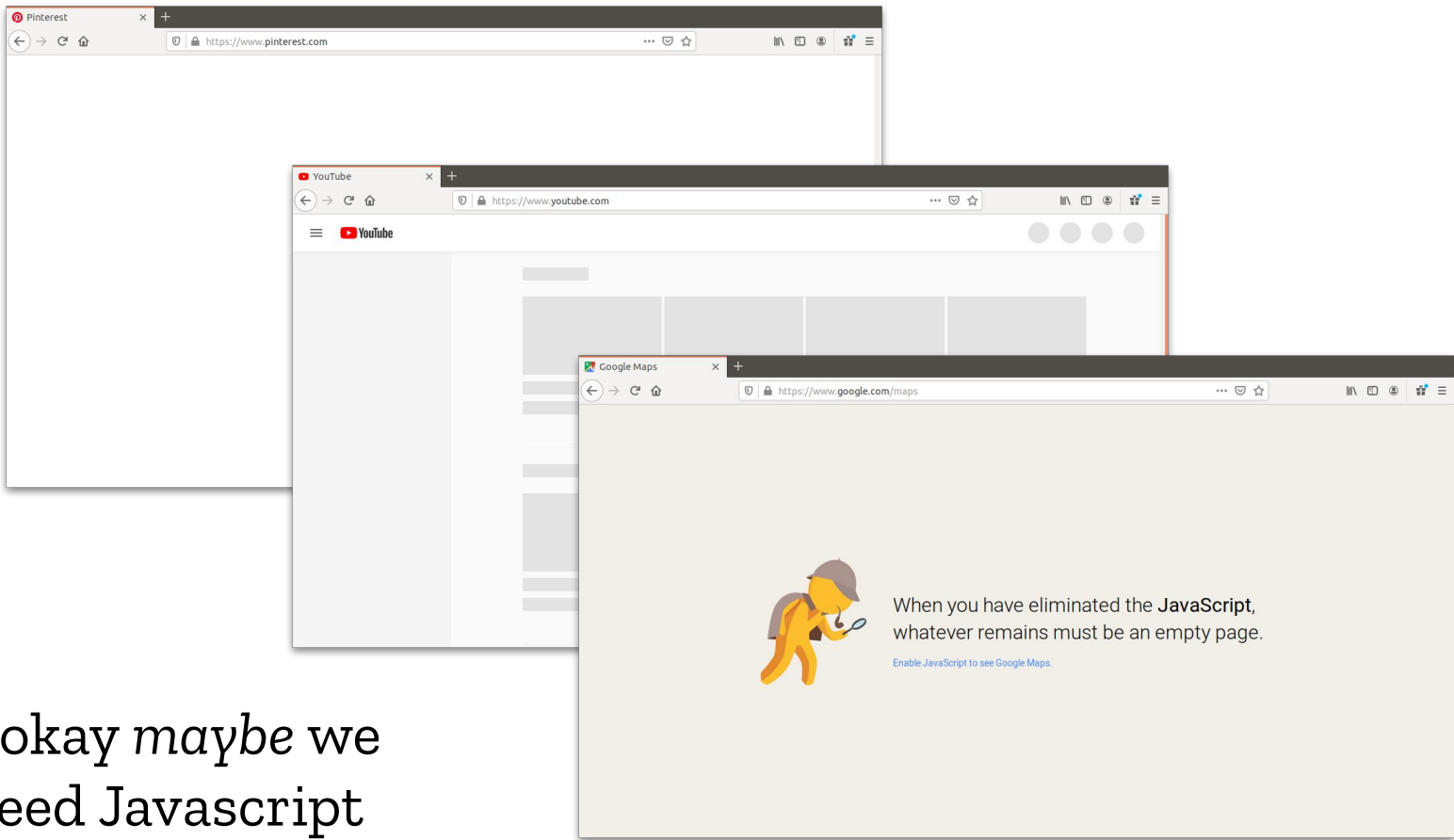
Our latest privacy experiment tested sites for an invisible form of online tracking that you can't easily avoid.



<https://www.washingtonpost.com/technology/2019/10/31/think-youre-anonymous-online-third-popular-websites-are-fingerprinting-you/>

Let's just block it all! Who needs Javascript or cookies?

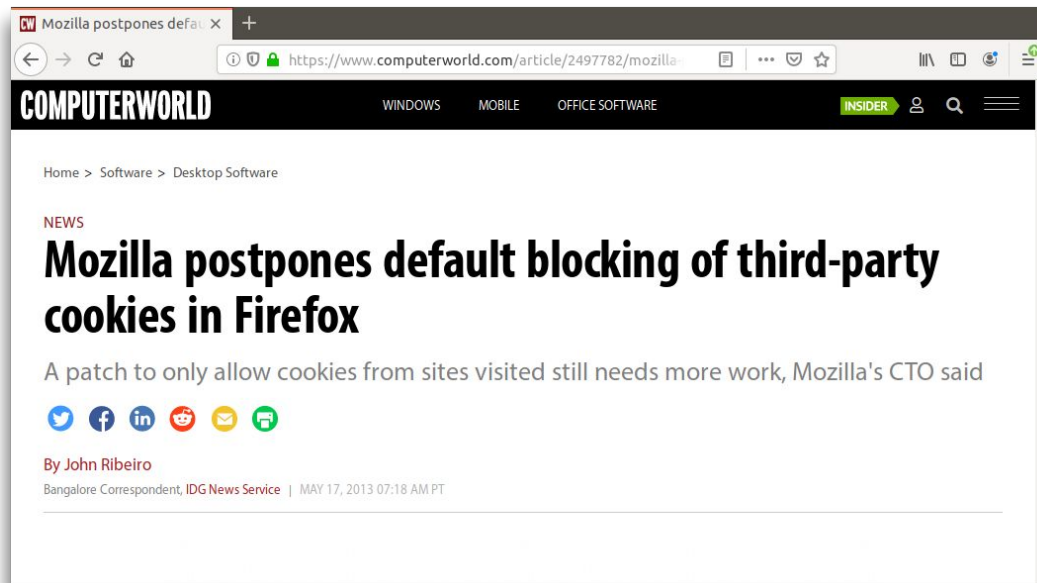




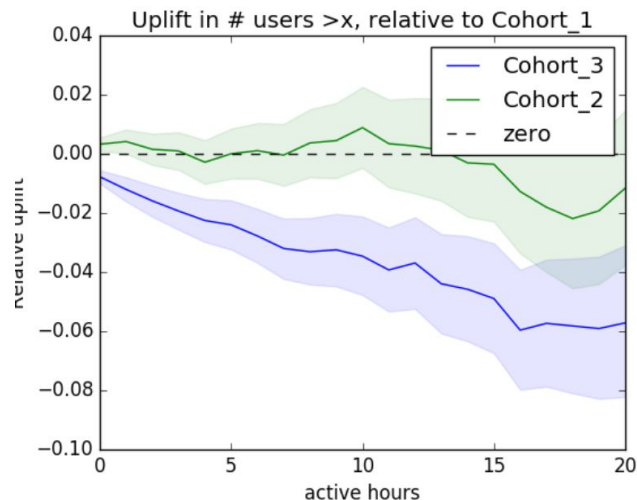
...okay *maybe* we
need Javascript

Can we at least block all third-party cookies?

We've tried that
back in 2013...



Can we at least block all third-party cookies?



Cohort_3: Block all third-party cookies
Cohort_2: Block cookies from trackers
Cohort_1: Control; No blocking

Compared to no blocking, users who had all third-party cookies blocked (over a 1 - 2 week study period):

- ~1% fewer active users overall
- 2 - 5% fewer users active over 10 hours

The drop was so significant, we ended this part of the study immediately.

https://mozilla.report/post/projects/cookie_restrictions.kp/

The core problem: one identifier, many uses



id=LgzcCiQvIbXtXjtkWrOZ

- Cross-site tracking
- Ad performance tracking
- Federated login
- First-party login
- Fraud detection
- Captcha / device reputation
- Analytics
- ... and many more

Our approach to anti-tracking...

1. No configuration necessary; **private by default**.
2. **Block the trackers**, not necessarily the tracking capability.
3. Don't break experiences **users care about**.

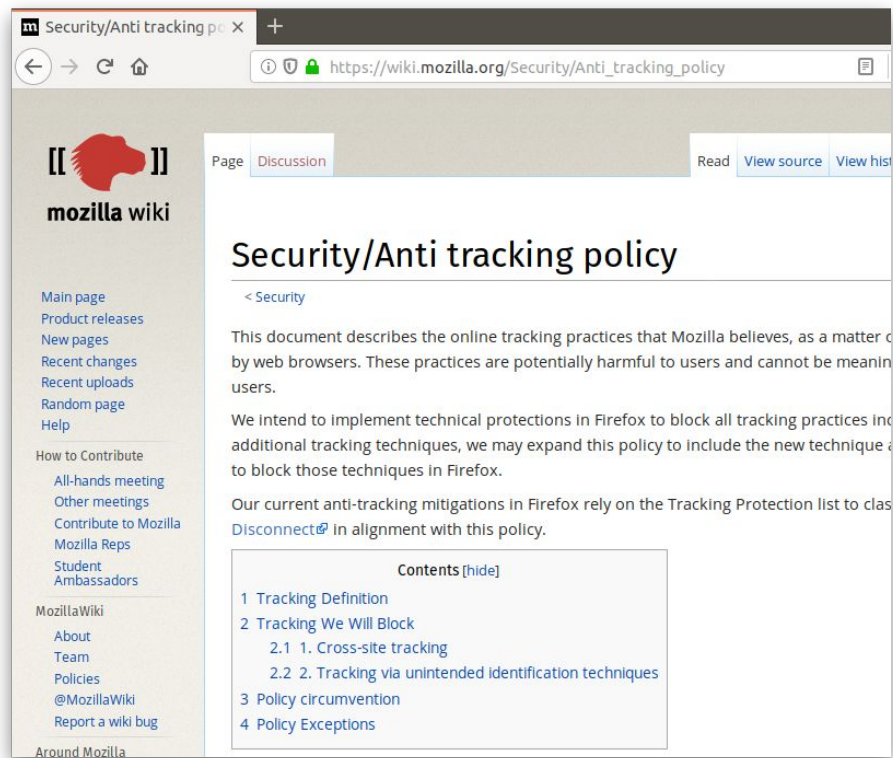
“[T]racking practices that Mozilla believes, as a matter of policy, **should be blocked** by default **by web browsers.**”

1. Cross-site tracking:

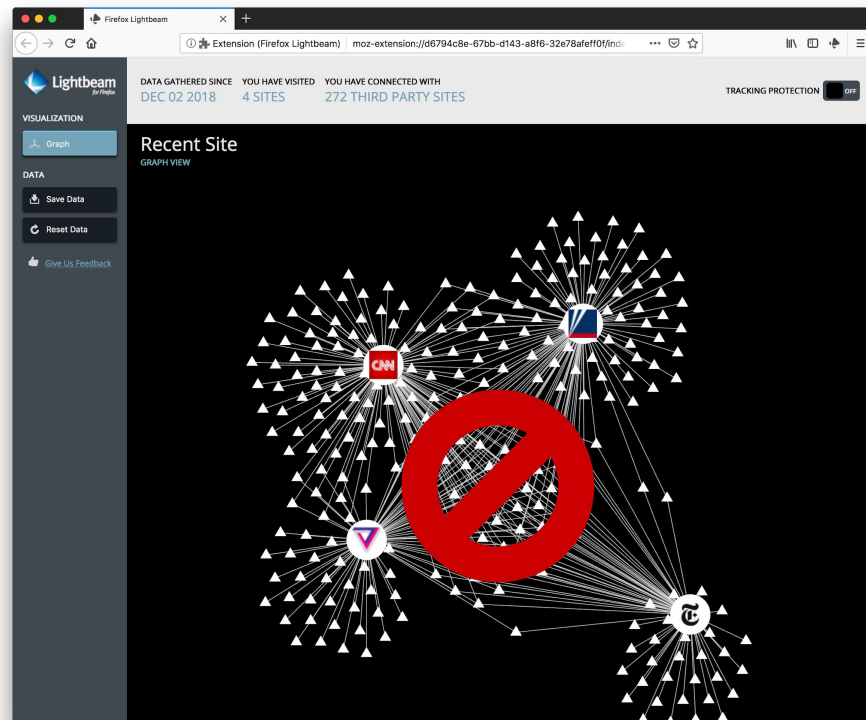
- a. Cookie-based
- b. URL parameter-based

2. Unintended identification techniques:

- a. Browser fingerprinting
- b. Supercookies



Our first step: Block only **tracking** cookies



Trackers are identified by Disconnect, based on a review of privacy policies.

It's more than just cookies...

We block all of this →

for all domains on the Disconnect Tracking Protection list

Cookies:

- Block `Cookie` request headers and ignore `Set-Cookie` response headers.
- Return an empty string for calls to `Document.cookie` and ignore requests to set cookies via `Document.cookie`.

DOM Storage:

- `localStorage`: `Window.localStorage` is `null`. Thus, attempts to read and write using this object will throw a `TypeError` exception.
- `sessionStorage`: read and write attempts are permitted.
- `IndexedDB`: read and write attempts throw a `SecurityError` exception.

Messaging and Workers:

- `BroadcastChannel`: attempts to create a new `BroadcastChannel` will throw a `SecurityError` exception.
- `SharedWorker`: attempts to create a new `SharedWorker` will throw a `SecurityError` exception.
- `Service Worker`: attempts to create a new `ServiceWorker` will throw a `SecurityError` exception.

DOM Cache:

- Calls to `CacheStorage` will always reject with a `SecurityError`.

Browser caches:

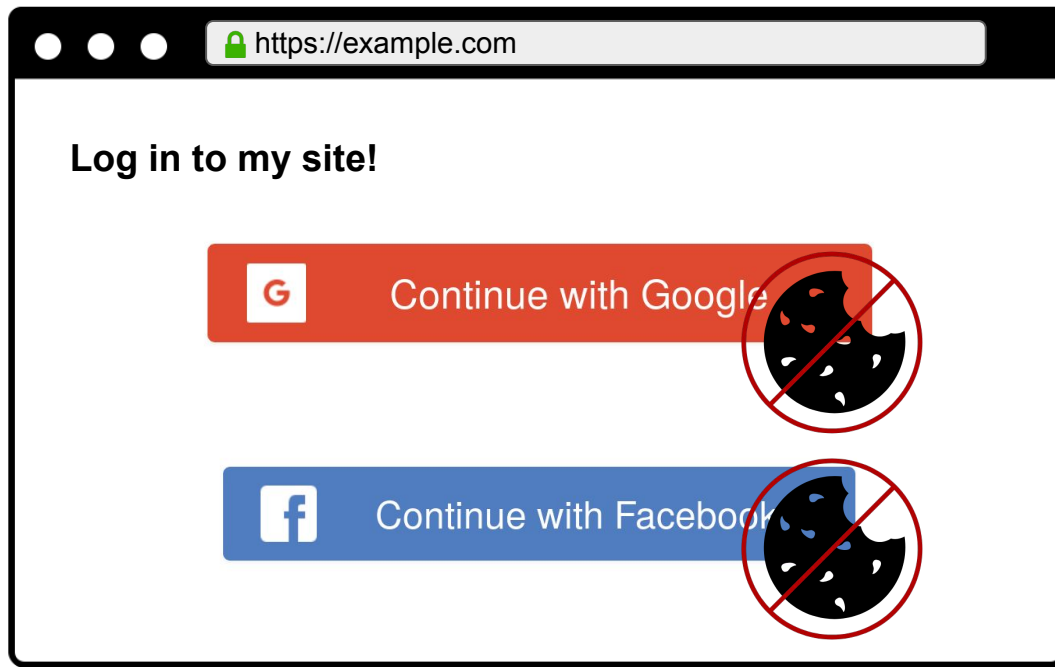
- The `HTTP cache` and the Image cache are partitioned for tracking resources, such that each top-level origin will have a separate partition and tracking resources on different top-level origins will be cached separate from each other.

Network connections:

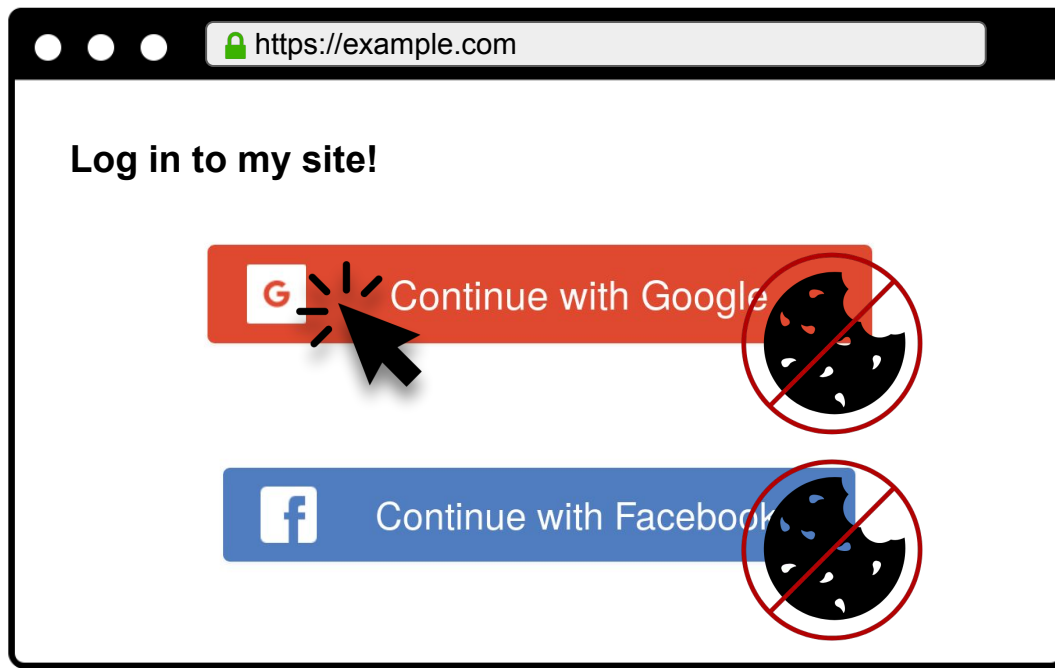
- `TLS sessions` will not be resumed using a session ticket when an HTTPS connection is made to an embedded third-party resource that is classified as a tracker.
- `HTTP connection reuse` by domains classified as trackers is limited to requests that occur under the same top-level origin. For example, a request for content from `tracker.example` on `news.example` will not reuse an HTTP connection with a request for content from `tracker.example` on `shopping.example` or with requests that occur when `tracker.example` is visited directly (i.e., as a first party).

https://developer.mozilla.org/en-US/docs/Mozilla/Firefox/Privacy/Storage_access_policy

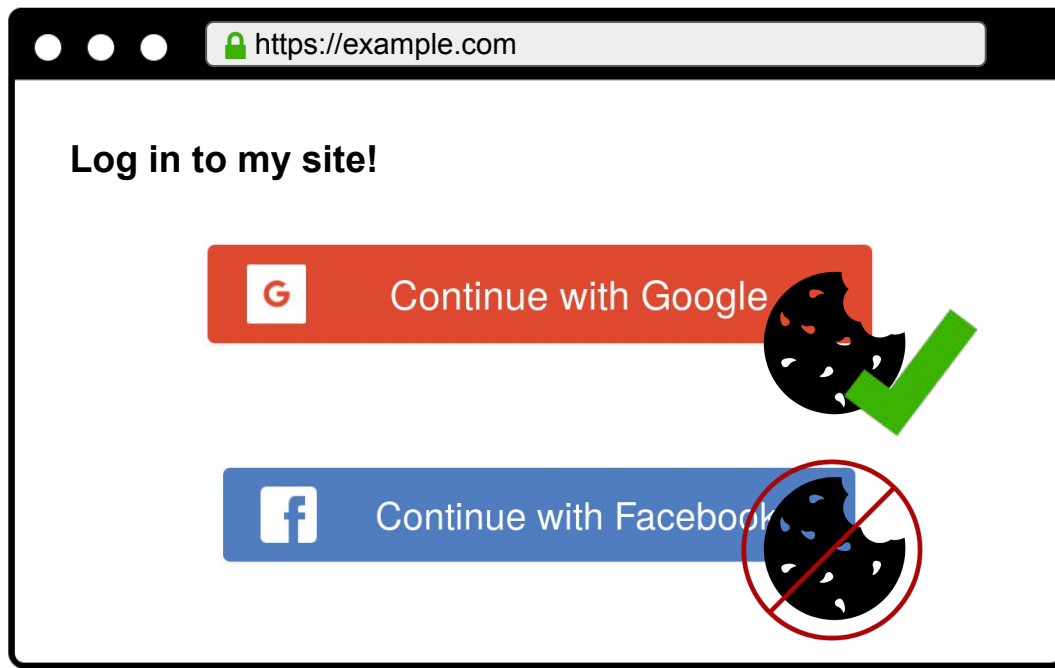
Some tracking is dual-use



Key insight: users interact with these services!



Cookies permitted after interaction, but **only on** example.com



https://developer.mozilla.org/en-US/docs/Mozilla/Firefox/Privacy/Storage_access_policy

Programmatic cookie access: the Storage Access API

Storage Access API methods

The storage API methods are implemented on the `Document` interface:

`Document.hasStorageAccess()`

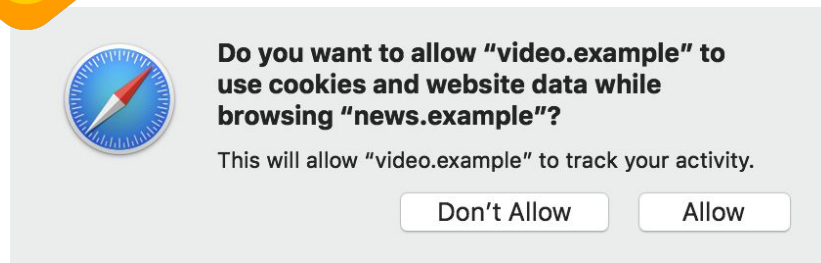
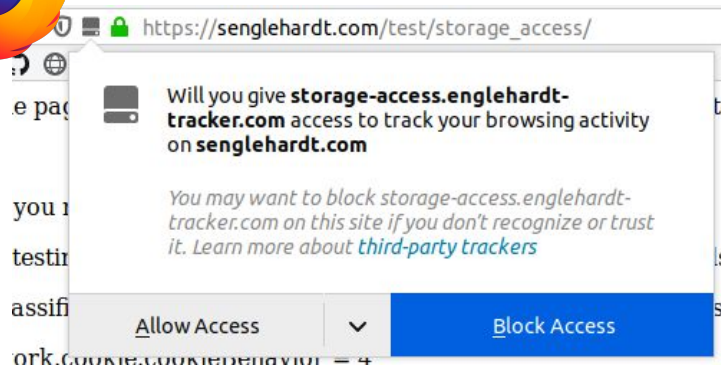
Returns a `Promise` that resolves with a boolean value indicating whether the document has access to its first-party storage.

`Document.requestStorageAccess()`

Returns a `Promise` that resolves if the access to first-party storage was granted, and rejects if access was denied.

https://developer.mozilla.org/en-US/docs/Web/API/Storage_Access_API

Programmatic cookie access: the Storage Access API



Notable differences in the prompting heuristics and scope of granted access:

- [https://developer.mozilla.org/en-US/docs/Web/API/Document/requestStorageAccess#Conditions for granting storage access](https://developer.mozilla.org/en-US/docs/Web/API/Document/requestStorageAccess#Conditions_for_granting_storage_access)
- [https://developer.mozilla.org/en-US/docs/Web/API/Storage_Access_API#Safari implementation differences](https://developer.mozilla.org/en-US/docs/Web/API/Storage_Access_API#Safari_implementation_differences)

We're seeing adoption of the Storage Access API

The diagram illustrates the adoption of the Storage Access API. It shows a Facebook comment interface on the left and an Amazon Pay dialog box on the right. A text box at the bottom, containing the code `calls Document.requestStorageAccess()`, has two arrows pointing to the 'Log In to Post' button on Facebook and the 'Continue' button on Amazon Pay. The Facebook interface includes a comment input area with a 'Hello!' placeholder, a 'Log In to Post' button, and a list of comments from Michael DiTraglia and Mildred R. Rosario. The Amazon Pay dialog box is titled 'Information' and 'Address book', showing a message about browser permissions for Amazon Pay and a 'Continue' button. The dialog also includes a 'Log out from Amazon Pay' link and a 'Save this information for next time' checkbox.

89 Comments Sort by Oldest

Hello!

☐ Also post on Facebook

Log In to Post

Michael DiTraglia
When you've lost Obama... https://www.realclearpolitics.com/.../obama_worried_that...
Like · Reply · 6w

Mildred R. Rosario

pay Address book
To complete your purchase, your browser needs your permission to access Amazon Pay.
Continue

Log out from Amazon Pay

☐ Save this information for next time

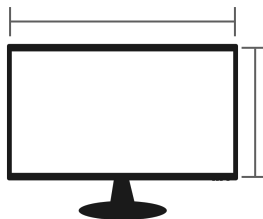
calls `Document.requestStorageAccess()`

The next step: browser fingerprinting

Locale



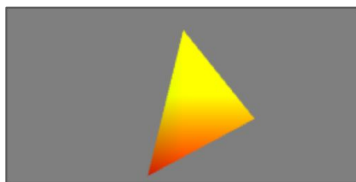
Screen Size



User Agent



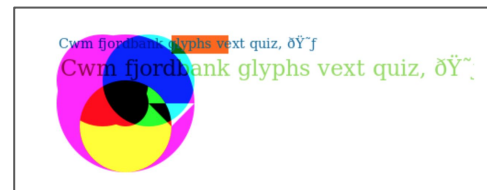
WebGL



Font probing

Times New Roman, Arial,
Open Sans, Courier New,
Georgia, Comic Sans, ...

HTML Canvas



Fingerprinting use is still growing...

2016

~1.6% of the Alexa
top 1M

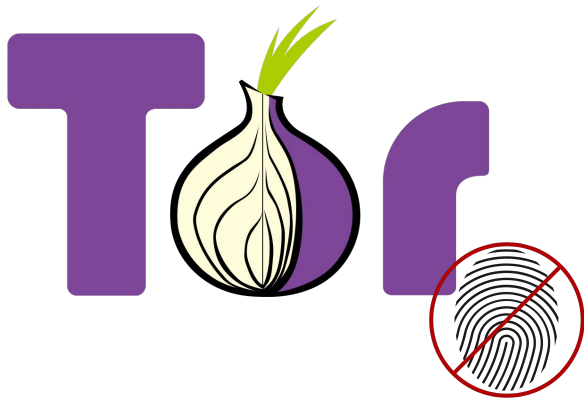
*Englehardt & Narayanan;
Online Tracking (CCS2016)*

2019

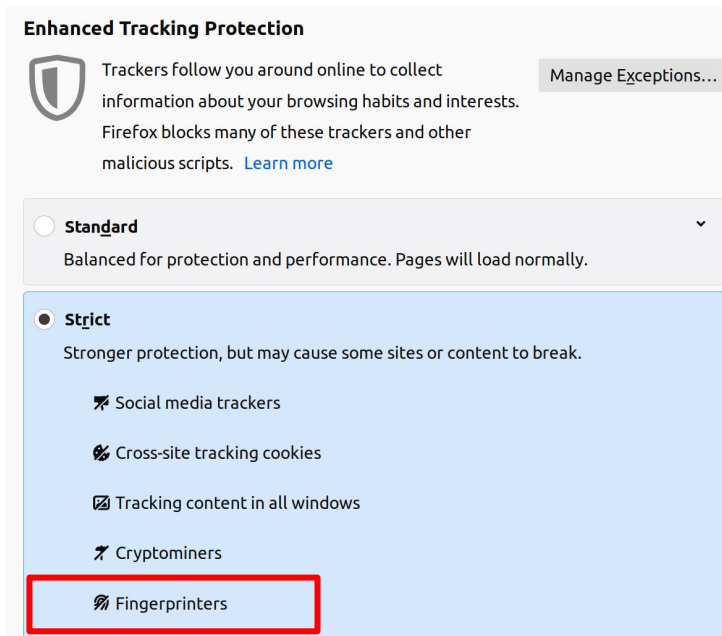
~3.8% of the
Alexa top 1M

*Unpublished OpenWPM
measurements from Feb 2019*

Two general approaches to anti-fingerprinting



Tor Browser's anti-fingerprinting in
Firefox behind
`privacy.resistFingerprinting`



Firefox 70

Bug 1507517 depends on 32 open bugs:

view as bug list | change several

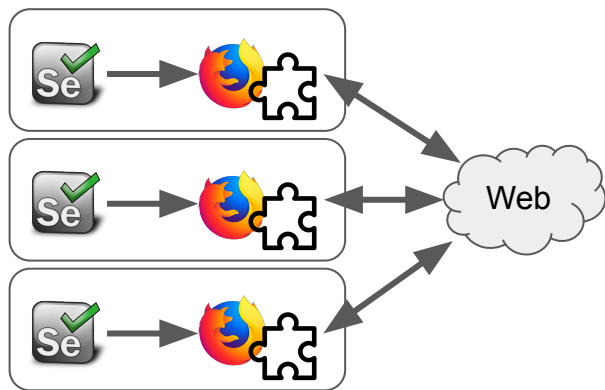
- 1507517: [META] Breakage from Fingerprinting Resistance
 - 1377744: privacy.resistFingerprinting's UTC timezone should not affect extensions
 - 1394448: Cannot install Addon with privacy.resistFingerprinting==true
 - 1404017: Pref for fingerprinting resistance in private browsing mode
 - 1394735: Enabling privacy.resistFingerprinting causes jank in jquery scrolling
 - 1399279: initial viewport too small for fullscreen WebApps with privacy.resistFingerprinting enabled
 - 1401493: Perform Fingerprint Comparison of Tor Browser and Firefox
 - 1414311: New window size is different than expected after changing screen dpi (with privacy.resistFingerprinting pref enabled)
 - 1428331: HiDPI and privacy.resistFingerprinting
 - 1554751: Consider to change the spoof value of window.devicePixelRatio
 - 1403099: game in <http://www.best.io/paper-io> has very bad performance due to anti-fingerprinting setting (needs higher resolution timer)
 - 1414311: New window size is different than expected after changing screen dpi (with privacy.resistFingerprinting pref enabled)
 - 1418537: Bad window height set when bookmarks toolbar is open with resistfingerprinting option
 - 1428331: HiDPI and privacy.resistFingerprinting
 - 1437266: Navigating back on youtube sometimes fails and restarts the current video with resistFingerprinting enabled
 - 1442863: Smooth scrolling implementations perform badly with resistFingerprinting's reduced timer precision
 - 1448423: browser.startup.blankWindow bugs when privacy.resistFingerprinting is enabled
 - 1448848: privacy.resistFingerprinting should not affect screen coordinates for extensions/content scripts
 - 1456378: privacy.resistFingerprinting breaks image cropping in Expensify
 - 1462115: privacy.resistFingerprinting affects the timezone displayed in native file picker dialogs
 - 1491343: Time is incorrect when the instance is opened via about:profiles in another profile with privacy.resistFingerprinting enabled
 - 1470828: privacy.resistFingerprinting breaks some shortcut keys
 - 1491343: Time is incorrect when the instance is opened via about:profiles in another profile with privacy.resistFingerprinting enabled
 - 1503872: reCAPTCHA v3 fails with Resist Fingerprinting Enabled
 - 1511941: privacy.resistfingerprinting performance API spoofing breaks vimeo.com
 - 1511982: chase.com login does not work when RFP is enabled
 - 1532859: privacy.resistFingerprinting makes Google Spreadsheet text blur
 - 1554751: Consider to change the spoof value of window.devicePixelRatio
 - 1533787: privacy.resistFingerprinting causes icons on some sites (including Gmail) to be blurry
 - 1554751: Consider to change the spoof value of window.devicePixelRatio
 - 1535565: [Wayland][resistFingerprinting] Maximized window remains garbled on startup until manually redrawn by switching windows
 - 1535568: [Wayland][resistFingerprinting] First maximized window dimensions are not being rounded down on startup
 - 1540308: privacy.resistFingerprinting set to true causes webpage to be white. Background image with z-index 5000 is not transparent.
 - 1554751: Consider to change the spoof value of window.devicePixelRatio
 - 1560816: privacy.resistFingerprinting should not return exact window dimensions as screen size
 - 1569561: wasm game doesn't run smoothly with privacy.resistFingerprinting enabled
 - 1573834: Uploading images on craigslist breaks with resistFingerprinting enabled
 - 1581492: [resistFingerprinting] Performance API spoofing prevents site from loading login scripts
 - 1589060: privacy.resistFingerprinting limits canvas webgl framerate to 10 fps

Changing APIs is hard...

- Image scaling problems from changing `devicePixelRatio`
- Image transparency issues
- Framerate and performance problems from timing changes

https://bugzilla.mozilla.org/show_bug.cgi?id=1507517

Our current approach: blocking fingerprinting scripts



Crawl the web with OpenWPM.
Detect fingerprinting scripts.

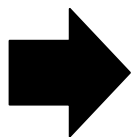
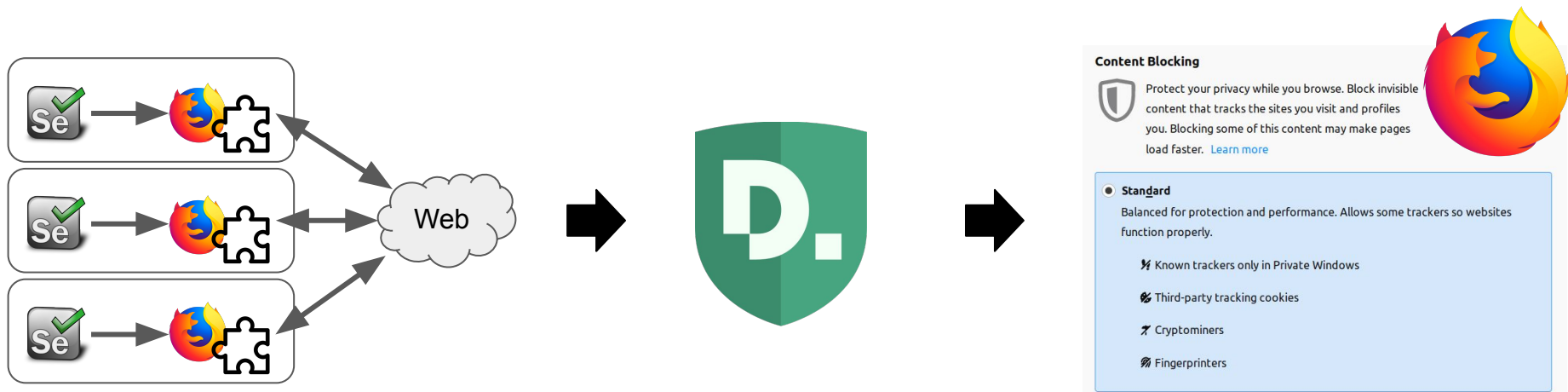


Share flagged scripts with
Disconnect, who does a review
to remove false positives.



Domains blocked in Firefox.
Eventually by default.

Our current approach: blocking fingerprinting scripts



138 verified fingerprinters
75% of the fingerprinting instances we've detected

Disconnect reviews candidate scripts

An example:

LeadsHub

This service has been classified as `Fingerprinting` for the following reasons:

Technical Review

Script: `http://cdn.ztsrv.com/js/0.5.0/ztag.js`

1. Script embeds or includes snippets of an open source fingerprinting library, [fingerprintjs2](#):

```
g = function() {
  if (!o()) return void 0;
  var t = document.createElement("canvas"),
      e = t.getContext("2d"),
      n = "http://valve.github.io";
  return e.textBaseline = "top", e.font = '14px "Arial"', e.textBaseline = "alphabetic", e.fillS
},
```

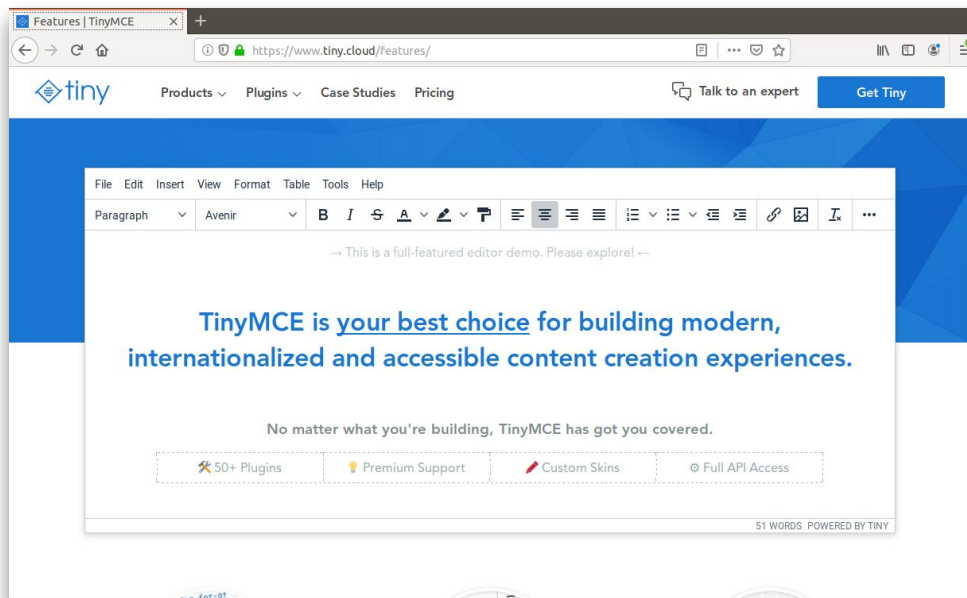
2. Sends computed fingerprint back to server

Request URL: `http://us-west-2-v2-t.ztsrv.com/1/i/REMOVED;za/p.gif`



<https://github.com/disconnectme/disconnect-tracking-protection/blob/master/descriptions.md>

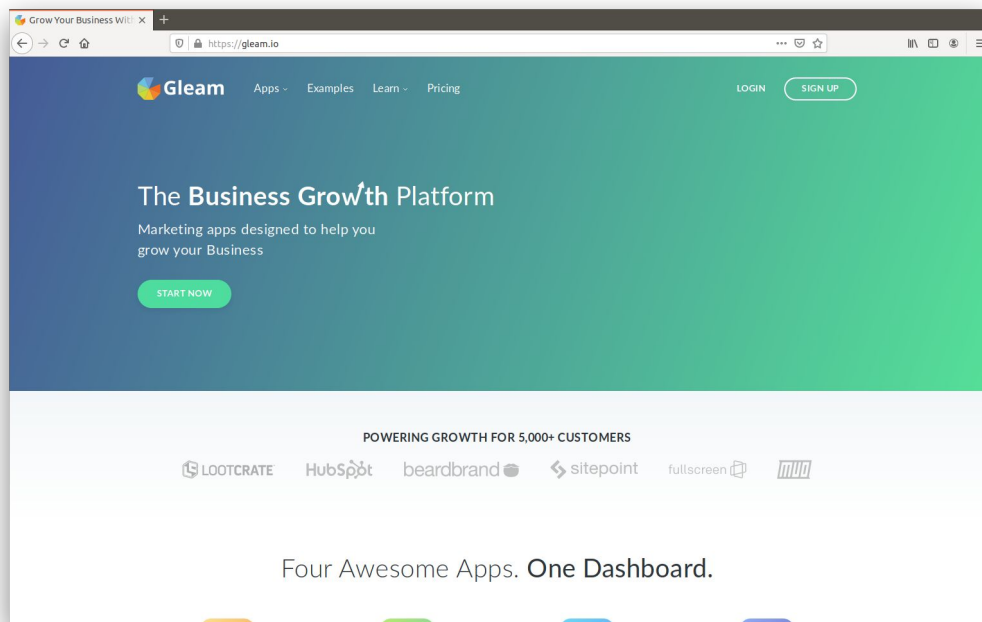
Initial success: changing practices



TinyMCE removed fingerprint2js from their HTML Editor as a result of our blocking

https://bugzilla.mozilla.org/show_bug.cgi?id=1544159

Initial success: changing practices

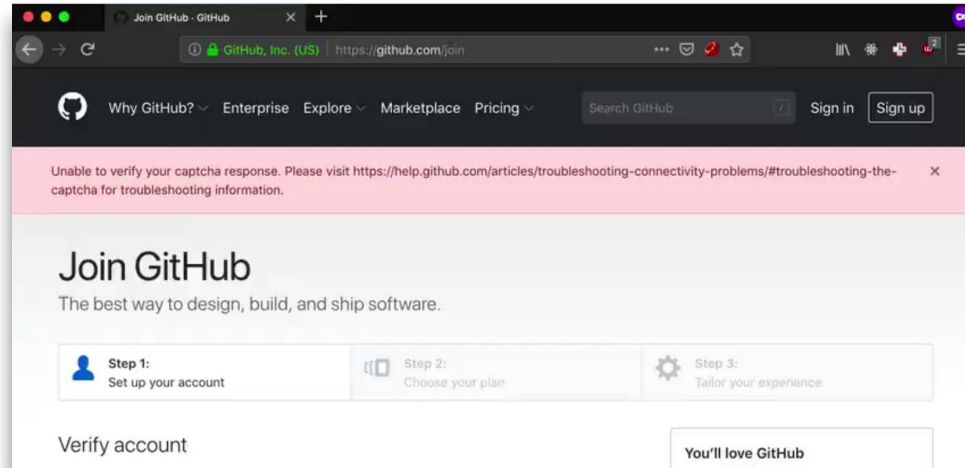


Glead moved fingerprinting script from `gleam.io` to `fraudjs.io` to avoid app breakage

https://bugzilla.mozilla.org/show_bug.cgi?id=1558658

The collage consists of three overlapping browser window screenshots. The leftmost window shows the Stripe website, featuring the Stripe logo, navigation links for 'Products' and 'Developers', a 'NEW' badge for 'Introducing Stripe Terminal', and the headline 'The new standard in online payment processing'. Below the headline, it states 'Stripe is the best software platform for running an internet business. We handle billions of dollars every year for forward-thinking businesses around the world.' and includes 'START NOW' and 'CONTACT SALES' buttons. The middle window shows the Sift Science website, with the headline 'ACCOUNT DEFENSE: Defend your users. Protect your brand.' and a sub-headline 'Put an end to fake accounts, and stop losing customers to account takeover.' A 'REQUEST A DEMO' button is visible. The rightmost window shows a screenshot of the Sift Science dashboard. It features a sidebar with navigation options like 'Reports', 'Overview', 'Accounts Report', 'Login Report', 'Logout Report', 'Payment Report', 'Device Report', 'Order Report', and 'Orders by Account Status'. The main content area displays an 'Accounts Report' for the time range 'September 2018 to October 2018'. It includes a bar chart showing 'Accounts Created' (Percentage of accounts blocked with total accounts created) and a table with columns for 'Date', 'Total Accounts Created', 'Total Accounts with Blocked Sessions', 'Accounts with Blocked Sessions', 'Accounts with Suspended Sessions', 'Total Accounts with Blocked Sessions', 'Accounts with Suspended Sessions', and 'Accounts with Suspended Sessions'. The table shows data for various dates, including 'Wed, Sep 12', 'Thu, Sep 13', 'Fri, Sep 14', 'Sat, Sep 15', and 'Sun, Sep 16'.

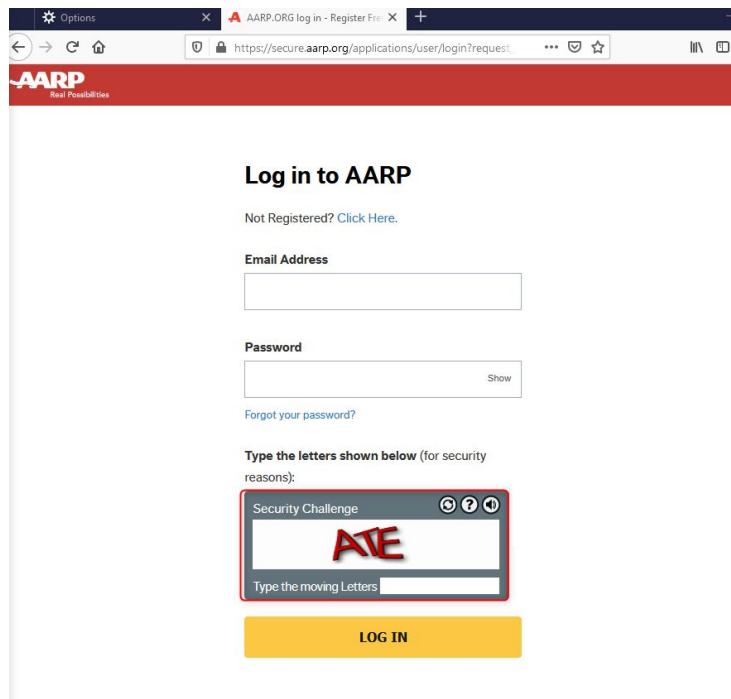
Blocking anti-fraud leads to major site breakage



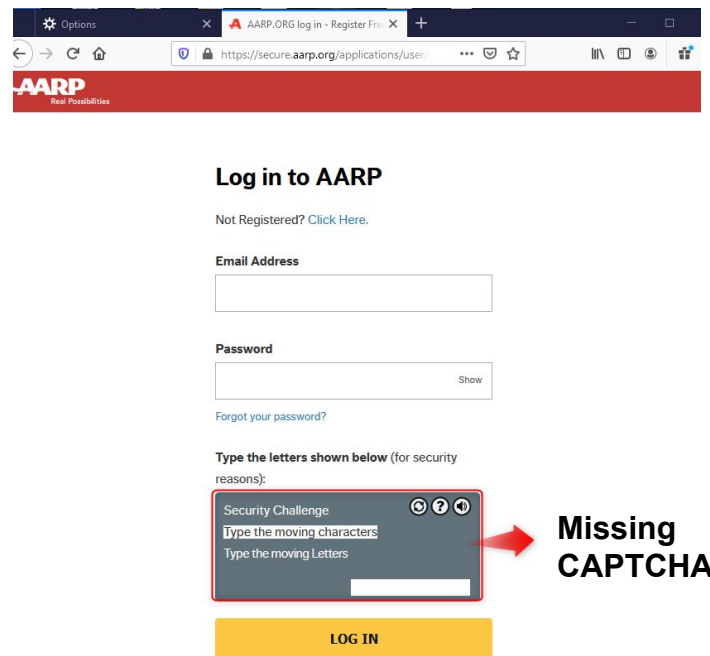
Github account creation broken when Arkose Labs captcha was blocked for fingerprinting

<https://twitter.com/richgilbank/status/1090269987888082944>

Blocking anti-fraud leads to major site breakage



No Blocking



Blocking Fingerprinters


Two possible solutions to safer anti-fraud?

Facebook Engineering

Open Source ▾ Platforms ▾ Infrastructure Systems ▾ Physical Infrastructure ▾ Video Engineering & AR/VR ▾ Artificial Intelligence ▾ Watch Videos

POSTED ON OCT 16, 2019 TO SECURITY

Fighting fraud using partially blind signatures



By Ben Savage Subodh Iyengar

Related Posts

- Aug 20, 2019
Redesigning our systems to provide more control over Off-Facebook activity
- Aug 15, 2019
Zoncolan: How Facebook uses static analysis to detect and prevent security issues

GitHub - WICG/trust-token-api

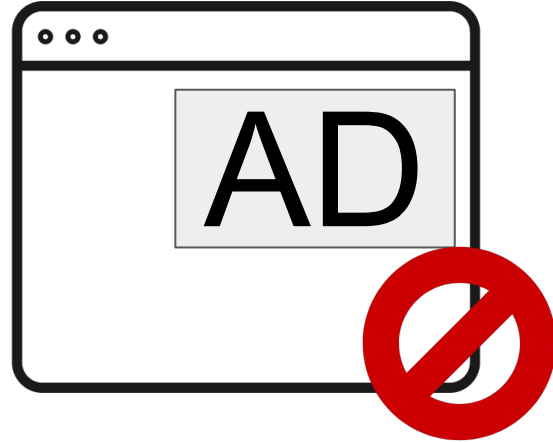
Trust Token API Explainer

This document is an explainer for a potential future web platform API that allows propagating trust across sites, using the [Privacy Pass](#) protocol as an underlying primitive.

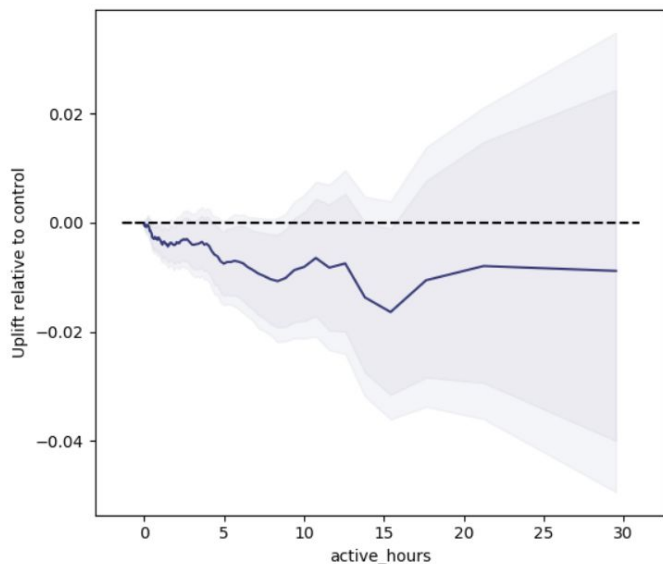
Table of Contents

- [Motivation](#)
- [Overview](#)
- [Potential API](#)
 - [Trust Token Issuance](#)
 - [Trust Token Redemption](#)
 - [Forwarding Redemption Attestation](#)
 - [Extension: Trust-Bound Keypair and Request Signing](#)
 - [Extension: Private Metadata](#)
- [Privacy Considerations](#)
 - [Cryptographic Property: Unlinkability](#)
 - [Key Consistency](#)
 - [Potential Attack: Side Channel Fingerprinting](#)
 - [Cross site Information Transfer](#)
 - [Mitigation: Dynamic Issuance / Redemption Limits](#)
 - [Mitigation: Allowed/Blocked Issuer Lists](#)
 - [Mitigation: Per-Site Issuer Limits](#)
 - [First Party Tracking Potential](#)
- [Security Considerations](#)
 - [Trust Token Exhaustion](#)
 - [Double-Spend Prevention](#)
- [Future Extensions](#)
 - [Publicly Verifiable Tokens](#)

Assume fraud is solved; can we block everything else?



Assume fraud is solved; can we block everything else?



... we tried that in a user study but we still saw 0.2% and 0.6% users leave Firefox because of it.

https://mozilla.report/post/projects/fingerprinting_protections.kp/

Why does blocking non-tracking fingerprinters
cause users to leave Firefox?

שני

Challenge: discovering sites broken by our protections

Approach	Problems
User Reports	<ul style="list-style-type: none">• Noisy• Unreliable
User Studies	<ul style="list-style-type: none">• Noisy• No clear way to measure
Manual QA	<ul style="list-style-type: none">• Limited scope• Expensive: 1 month of full-time work per 1,000 sites
Automated Crawls	<ul style="list-style-type: none">• Limited scope• No clear way to measure

Mozilla Research Grants 2019H2

How can we automate the process of discovering broken sites?

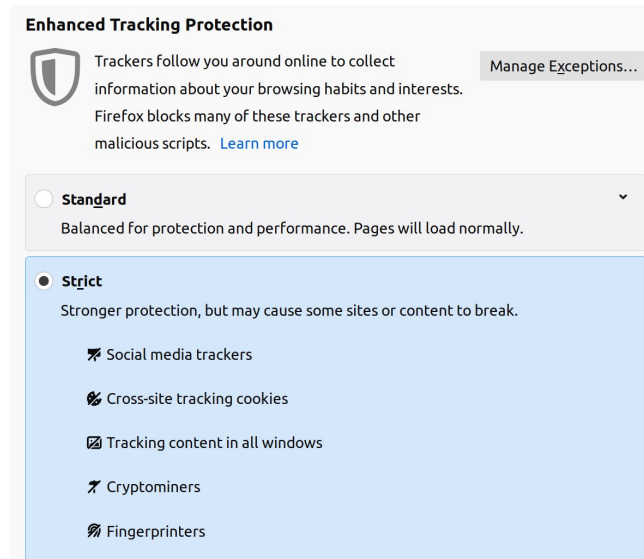
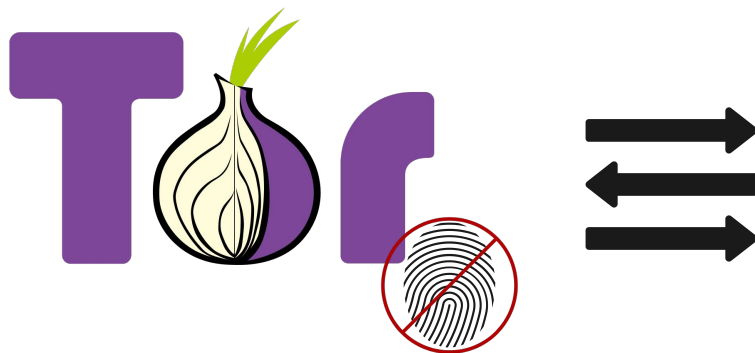
Amount: \$40,000

Deadline

Friday, November 22nd at 2:22pm
Pacific Daylight Time (PDT)

<https://mozilla-research.forms.fm/mozilla-research-grants-2019h2/forms/7376>

A possible step forward for anti-fingerprinting?



Per-frame fingerprinting resistance based on a blocklist

https://bugzilla.mozilla.org/show_bug.cgi?id=1531873

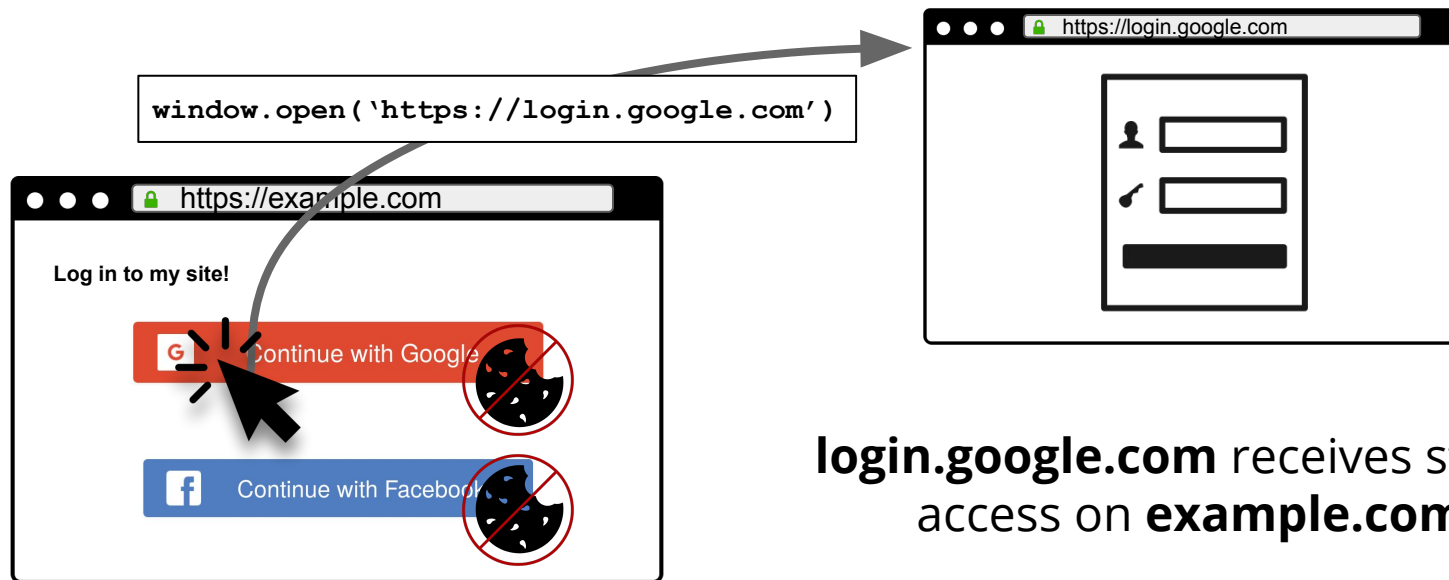
My asks for you:

1. Help us find technical alternatives to a global identifier for:
 - a. Federated login
 - b. Anti-fraud / device reputation
 - c. Advertisement attribution / measurement
2. Find violations of our anti-tracking policies
 - a. Name and shame
 - b. We can update our blocks
3. Help us explore ways to better discover broken sites
 - a. Apply to our grant!

THANK YOU!

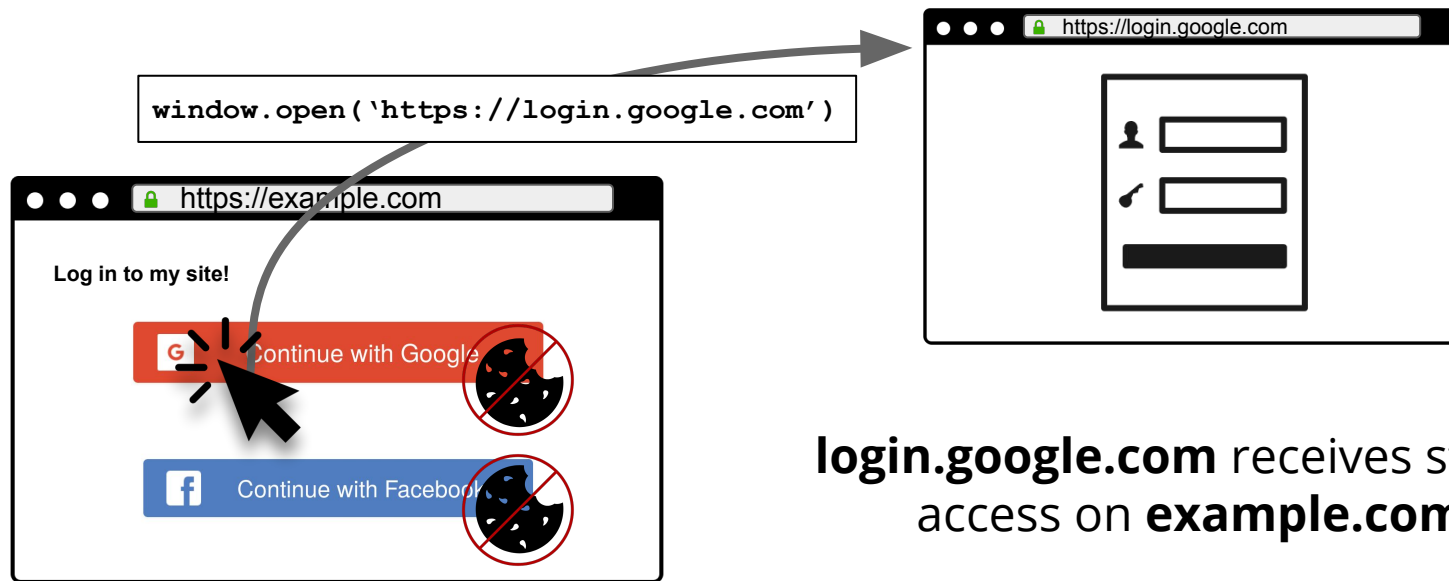
BACKUP SLIDES

Interaction allows interactive embeds to work



login.google.com receives storage access on **example.com***

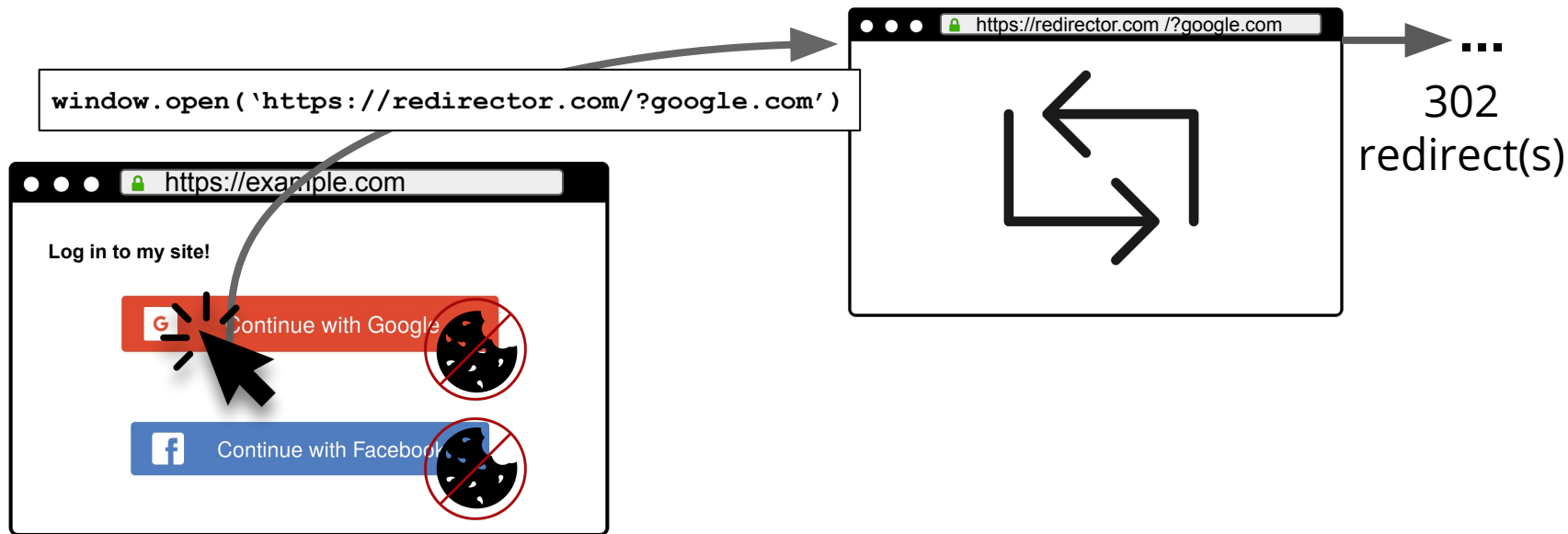
Interaction allows interactive embeds to work



** Provided it meets some additional requirements.*

See: https://developer.mozilla.org/en-US/docs/Mozilla/Firefox/Privacy/Storage_access_policy

We also handle redirects...



See: https://developer.mozilla.org/en-US/docs/Mozilla/Firefox/Privacy/Storage_access_policy

A workaround: Cross-site tracking with query strings



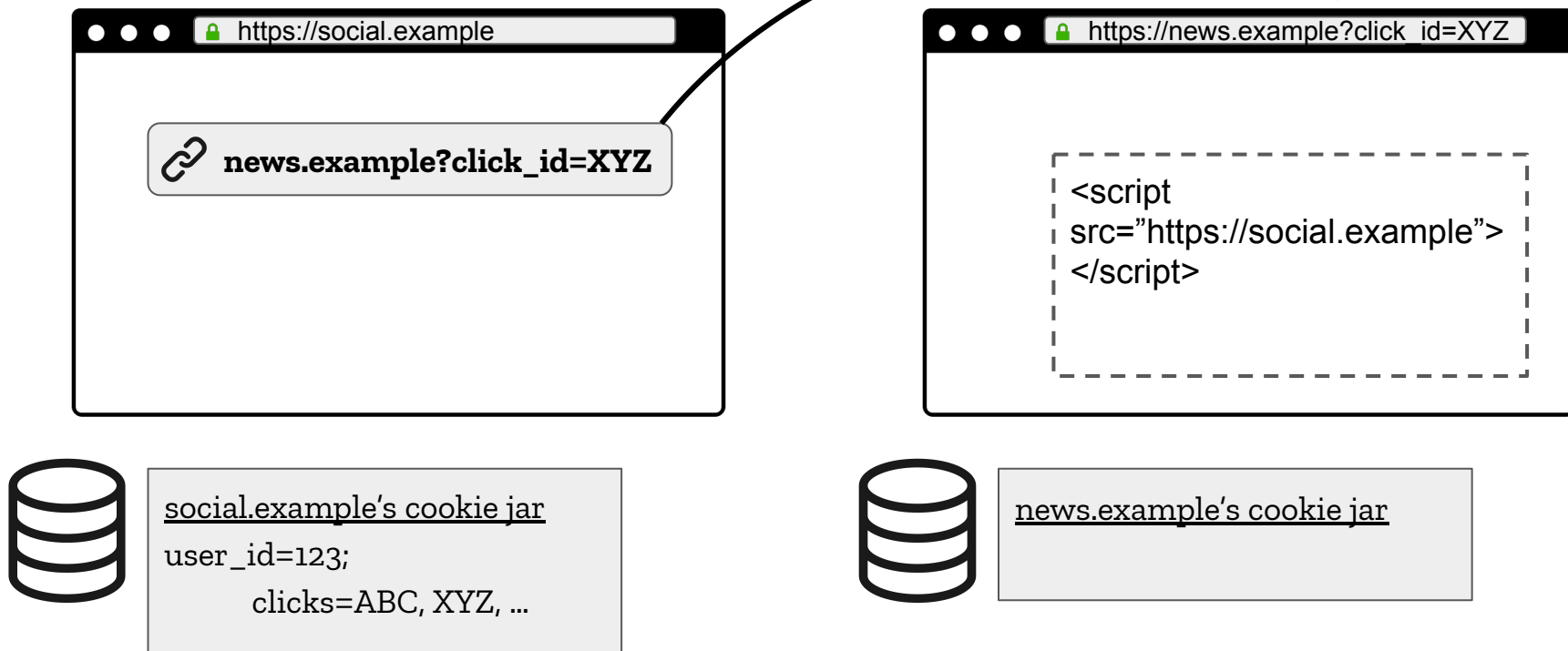
social.example's cookie jar
user_id=123;
clicks=ABC, ...



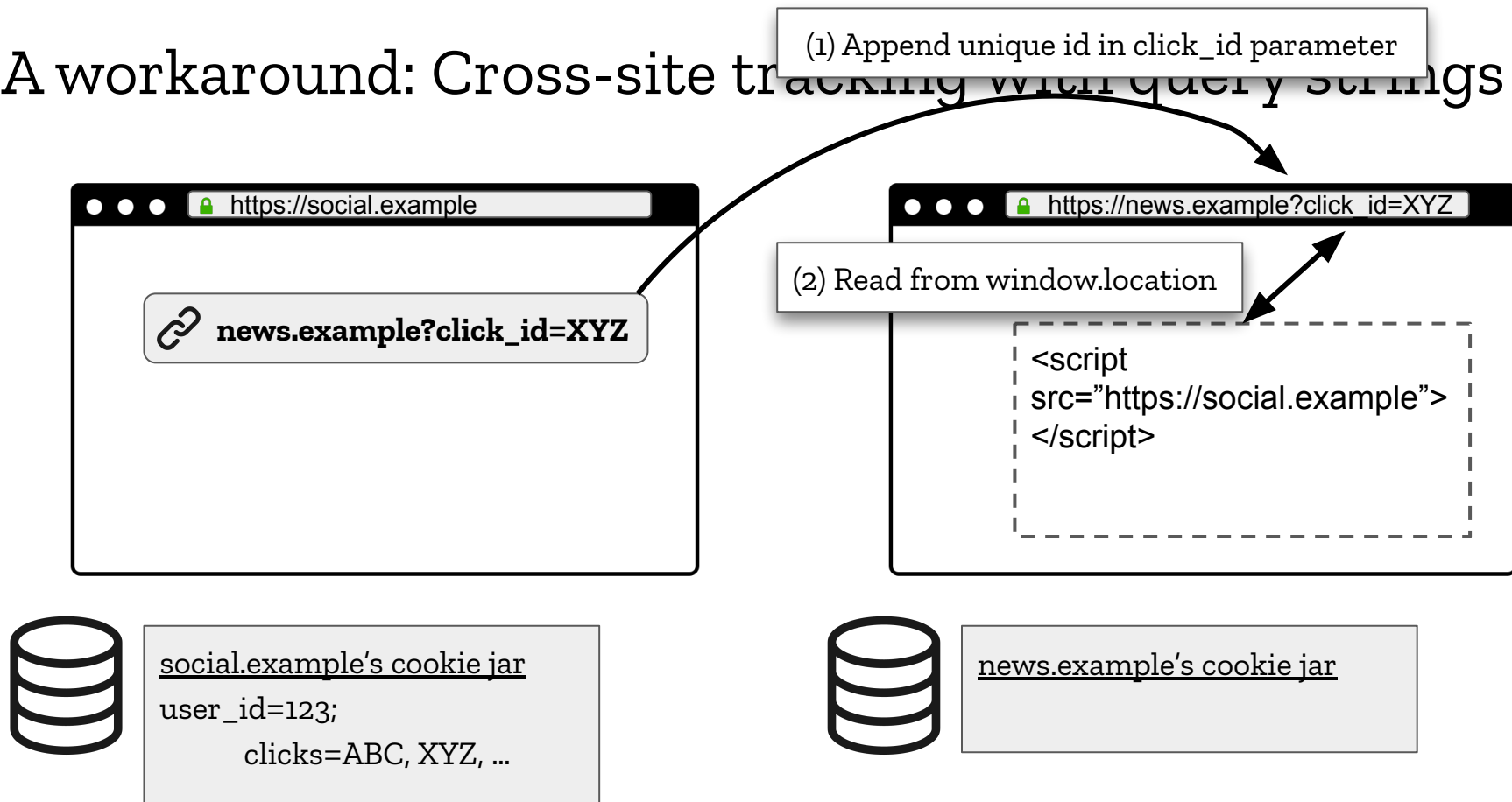
news.example's cookie jar

A workaround: Cross-site tracking with query strings

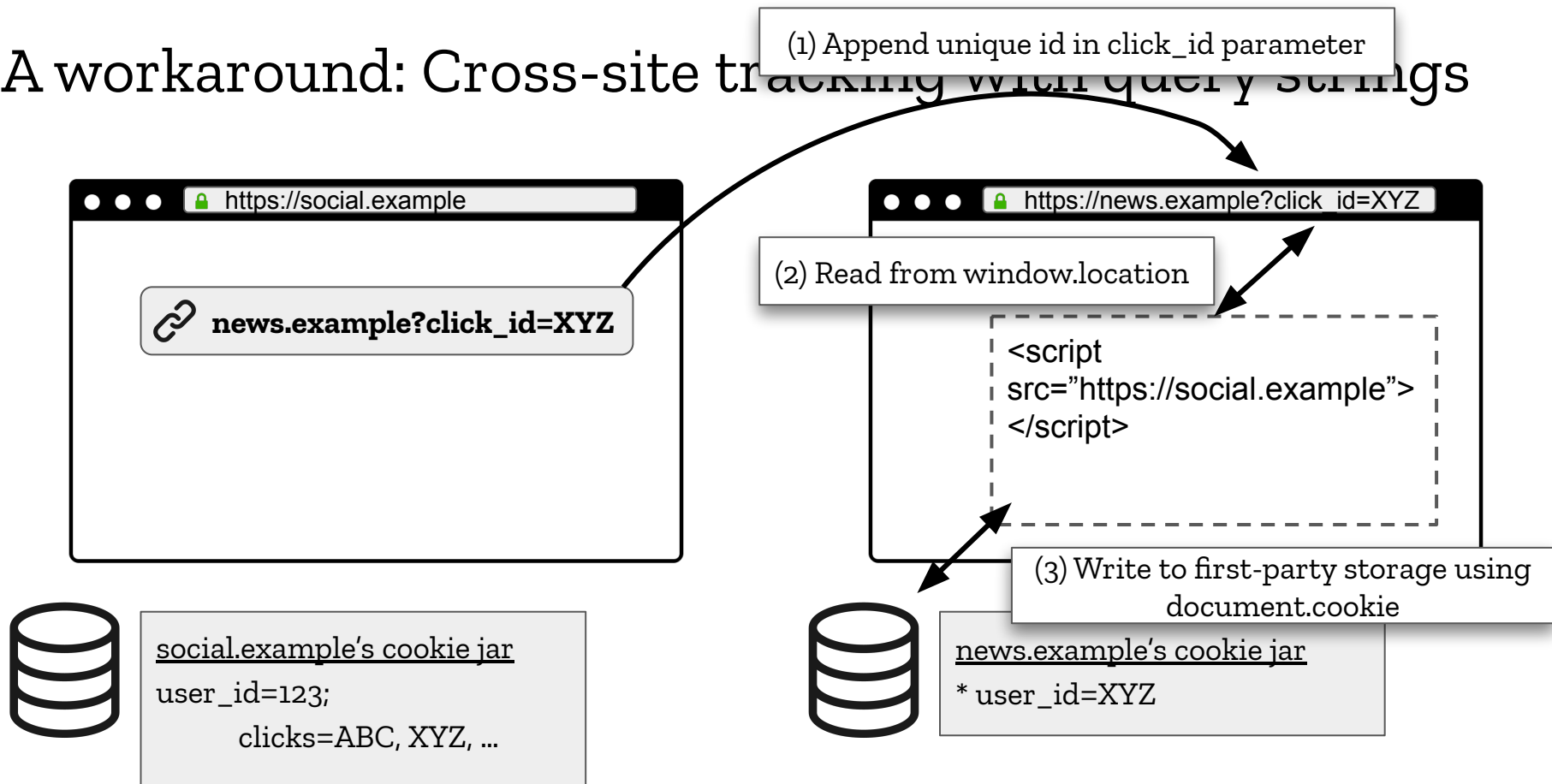
(1) Append unique id in click_id parameter



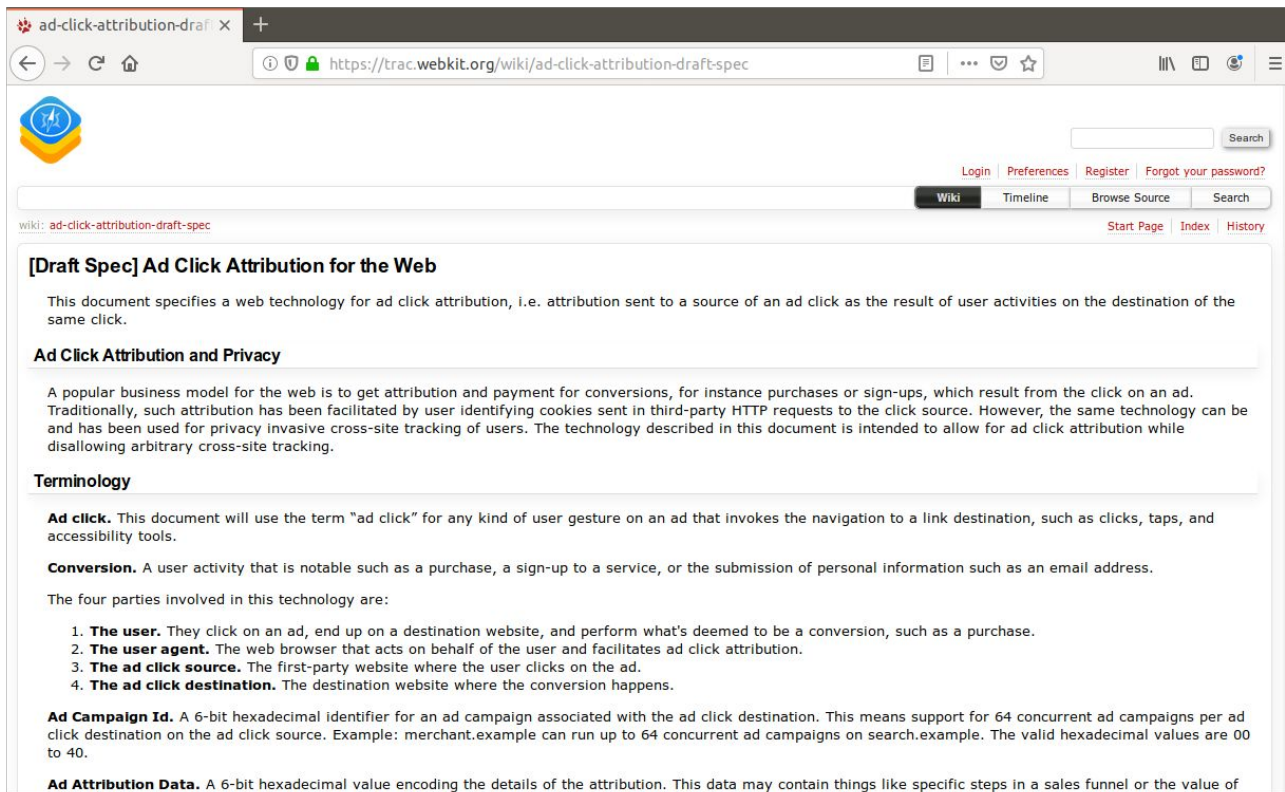
A workaround: Cross-site tracking with query strings



A workaround: Cross-site tracking with query strings



Request: a safer way to do ad measurement



The screenshot shows a web browser window with the address bar displaying `https://trac.webkit.org/wiki/ad-click-attribution-draft-spec`. The page is a draft specification document titled "[Draft Spec] Ad Click Attribution for the Web". The document content includes:

- [Draft Spec] Ad Click Attribution for the Web**

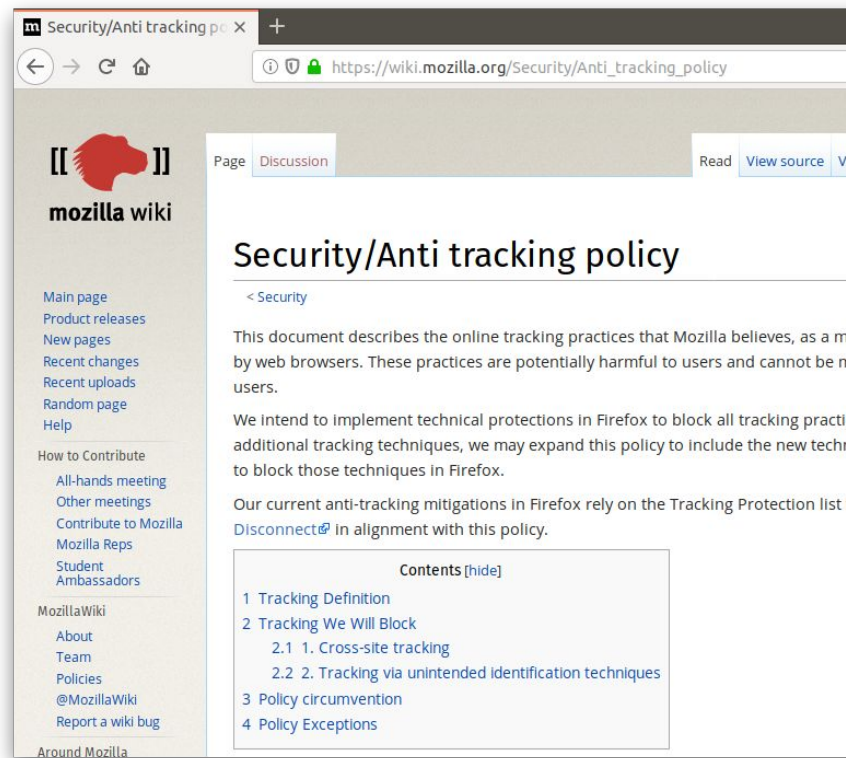
This document specifies a web technology for ad click attribution, i.e. attribution sent to a source of an ad click as the result of user activities on the destination of the same click.
- Ad Click Attribution and Privacy**

A popular business model for the web is to get attribution and payment for conversions, for instance purchases or sign-ups, which result from the click on an ad. Traditionally, such attribution has been facilitated by user identifying cookies sent in third-party HTTP requests to the click source. However, the same technology can be and has been used for privacy invasive cross-site tracking of users. The technology described in this document is intended to allow for ad click attribution while disallowing arbitrary cross-site tracking.
- Terminology**
 - Ad click.** This document will use the term "ad click" for any kind of user gesture on an ad that invokes the navigation to a link destination, such as clicks, taps, and accessibility tools.
 - Conversion.** A user activity that is notable such as a purchase, a sign-up to a service, or the submission of personal information such as an email address.
- The four parties involved in this technology are:
 1. **The user.** They click on an ad, end up on a destination website, and perform what's deemed to be a conversion, such as a purchase.
 2. **The user agent.** The web browser that acts on behalf of the user and facilitates ad click attribution.
 3. **The ad click source.** The first-party website where the user clicks on the ad.
 4. **The ad click destination.** The destination website where the conversion happens.
- Ad Campaign Id.** A 6-bit hexadecimal identifier for an ad campaign associated with the ad click destination. This means support for 64 concurrent ad campaigns per ad click destination on the ad click source. Example: merchant.example can run up to 64 concurrent ad campaigns on search.example. The valid hexadecimal values are 00 to 40.
- Ad Attribution Data.** A 6-bit hexadecimal value encoding the details of the attribution. This data may contain things like specific steps in a sales funnel or the value of

Challenge: Fingerprinting for anti-fraud

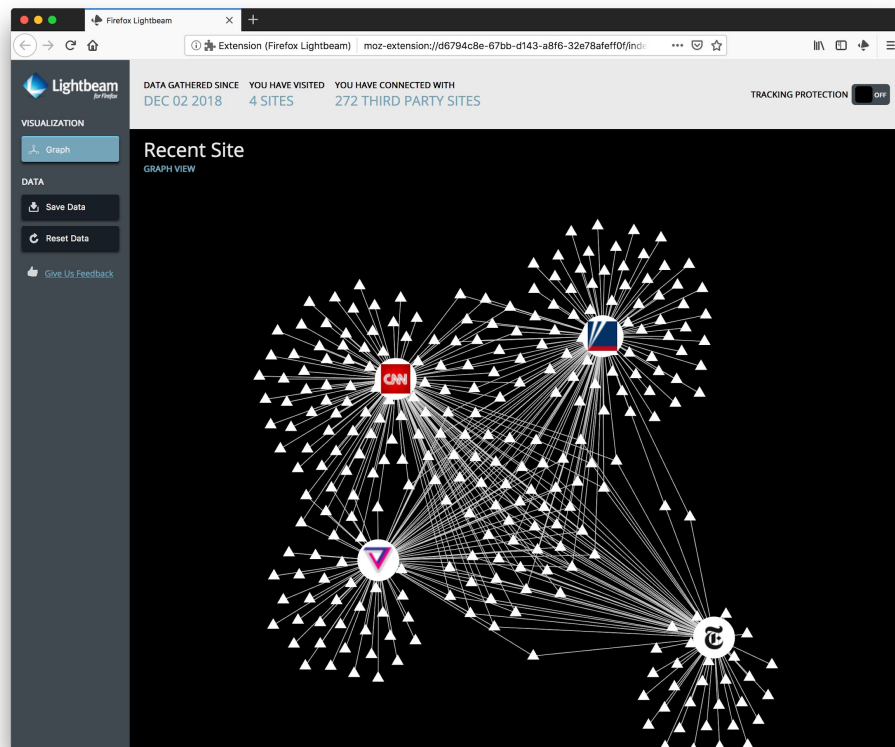
Policy exceptions:

1. Improving client authentication
2. Preventing the creation of fraudulent accounts
3. Preventing the completion of fraudulent purchases.



The web needs **default-on** tracking protection ...

4 news sites
272 third parties



The web needs **default-on** tracking protection ...
... and not just from third-party cookies



Browser state



IP Address + Device
Properties



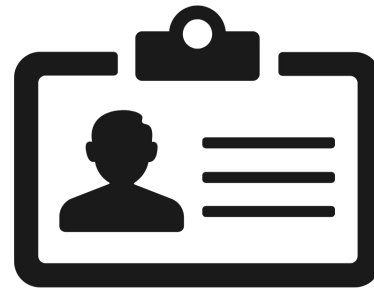
Identity



Browser state



IP Address + Device
Properties



Identity

**Tracking vectors completely
within browser's control**

The tracking landscape



Browser state



IP Address + Device
Properties



Identity

The tracking landscape



Browser state



IP Address + Device
Properties

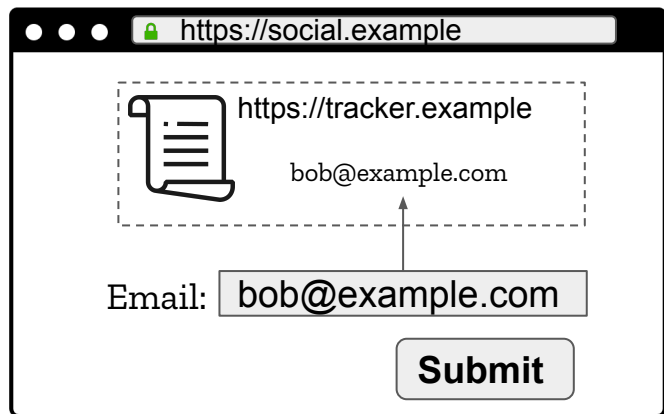


Identity

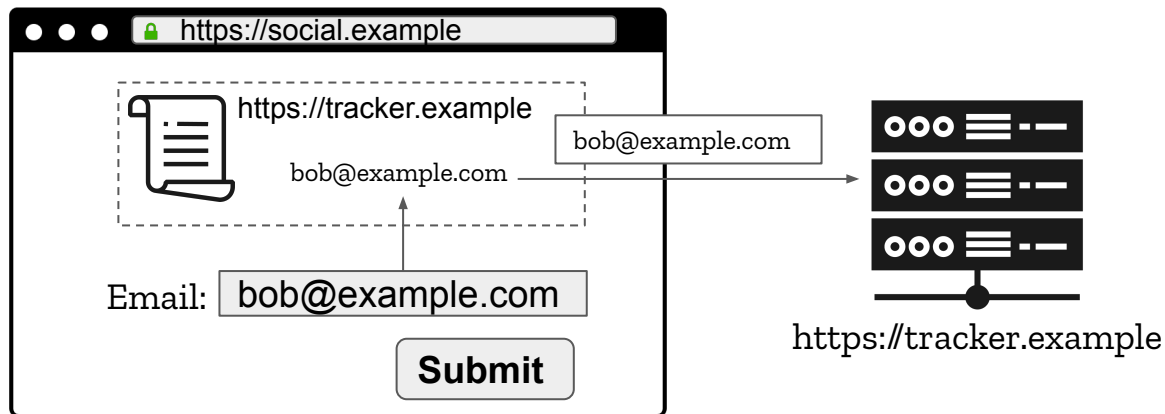
Scripts can collect PII for tracking



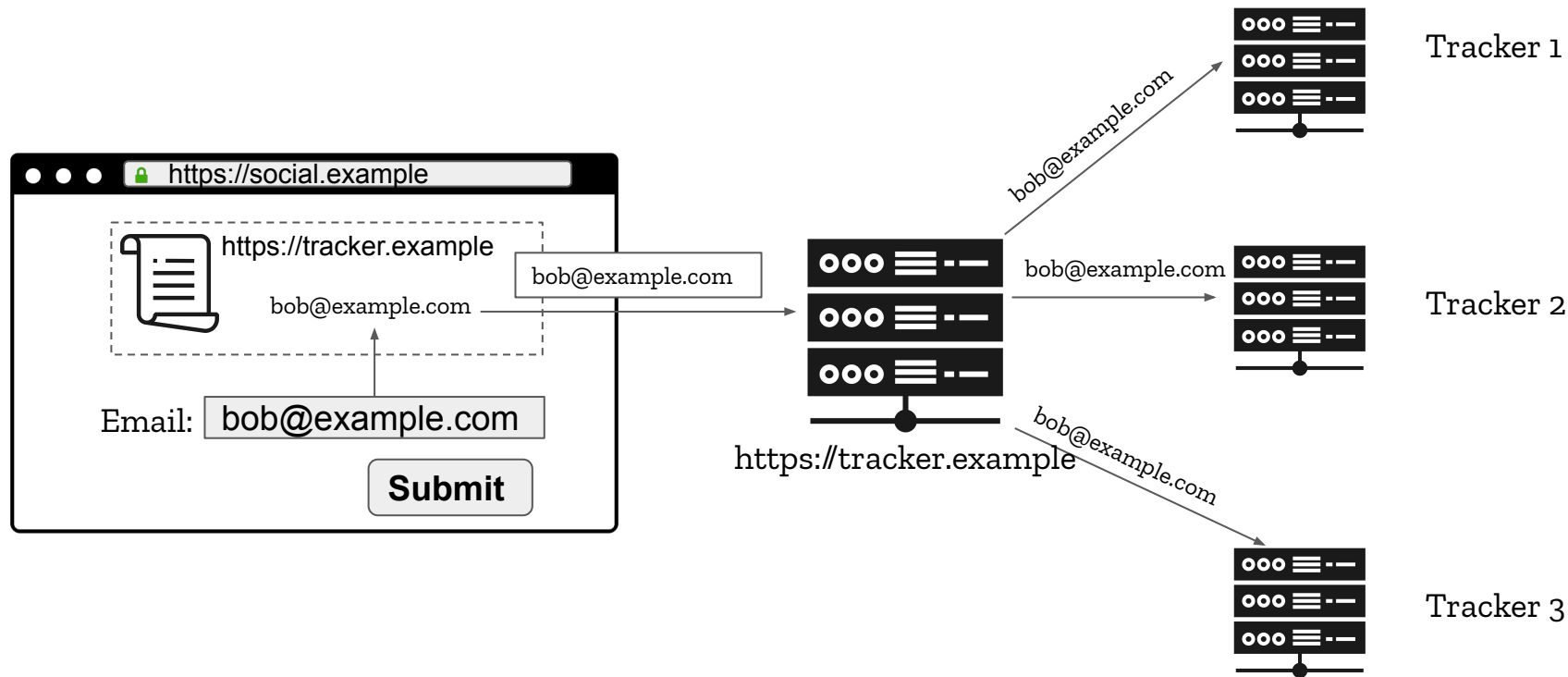
Scripts can collect PII for tracking



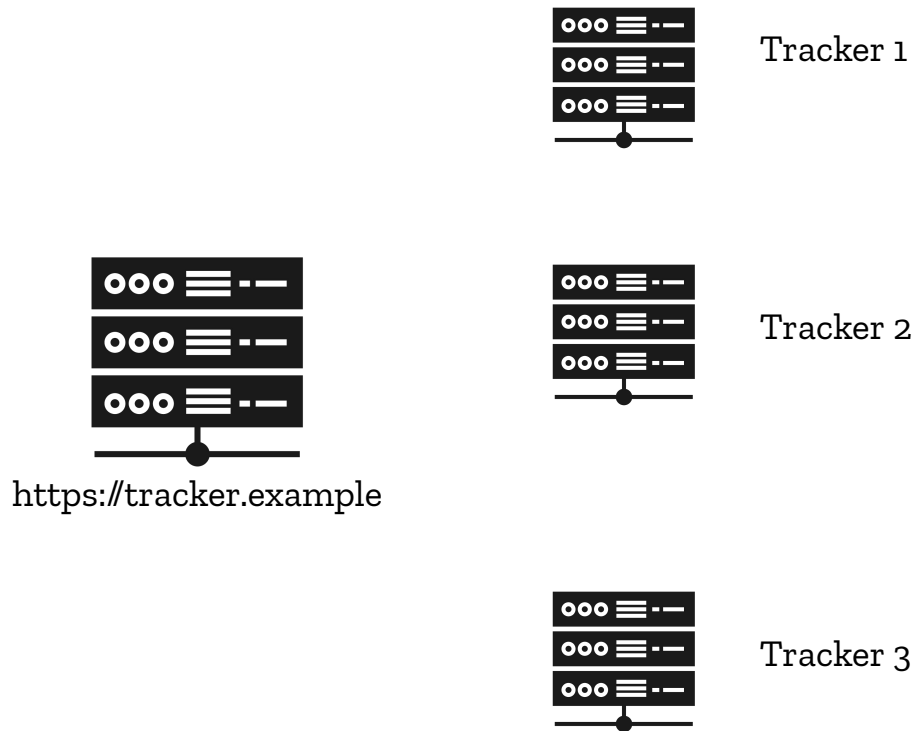
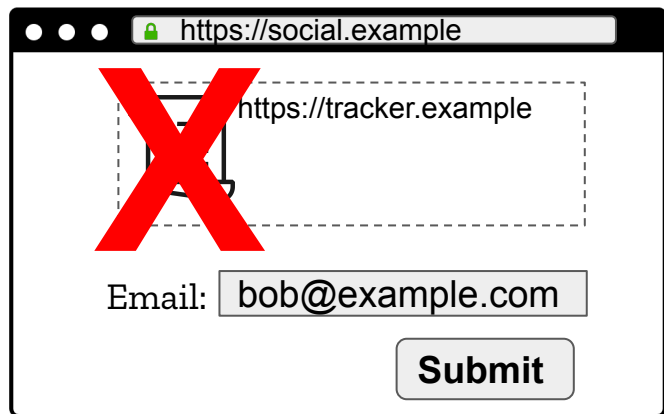
Scripts can collect PII for tracking



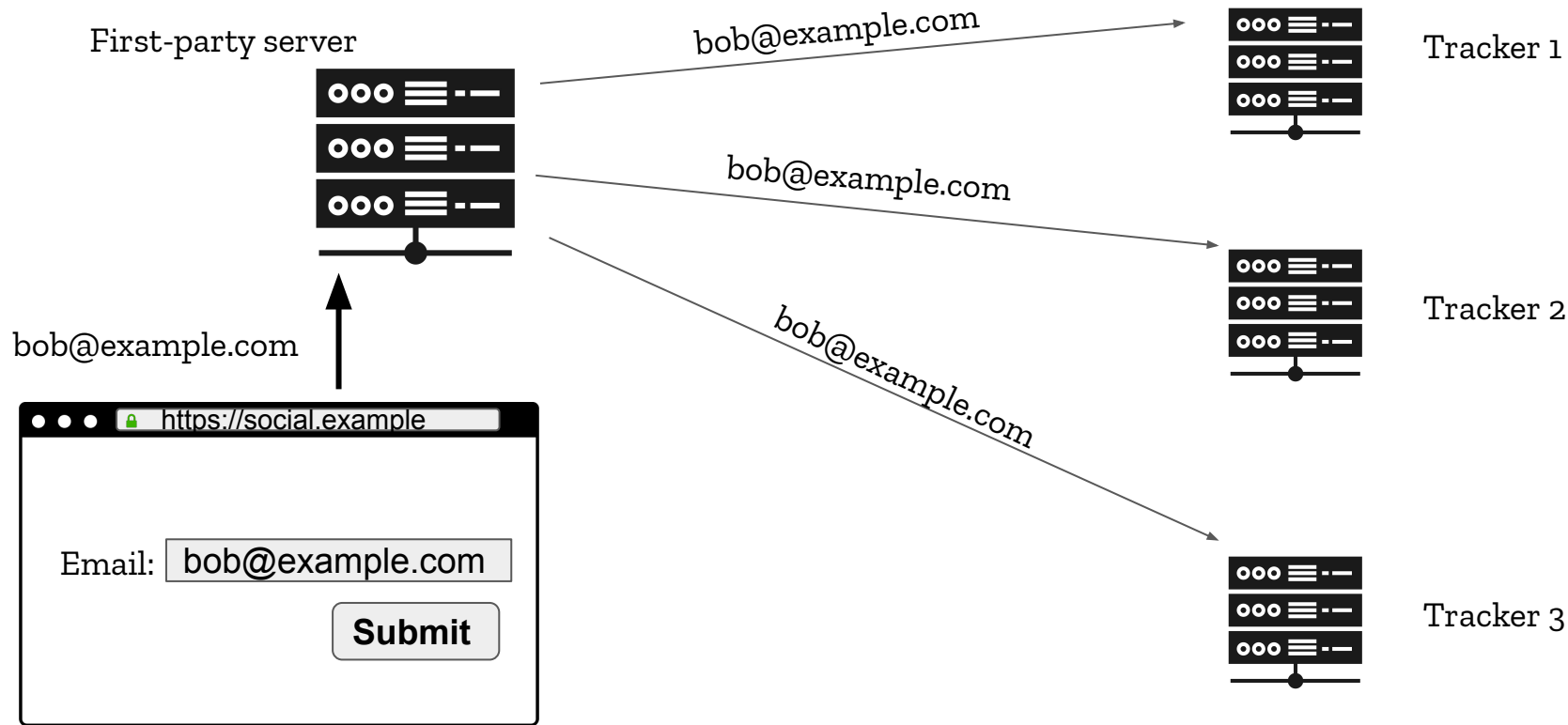
Scripts can collect PII for tracking



We can block scripts that scrape PII



In-browser protection options are limited...



The tracking landscape



Browser state



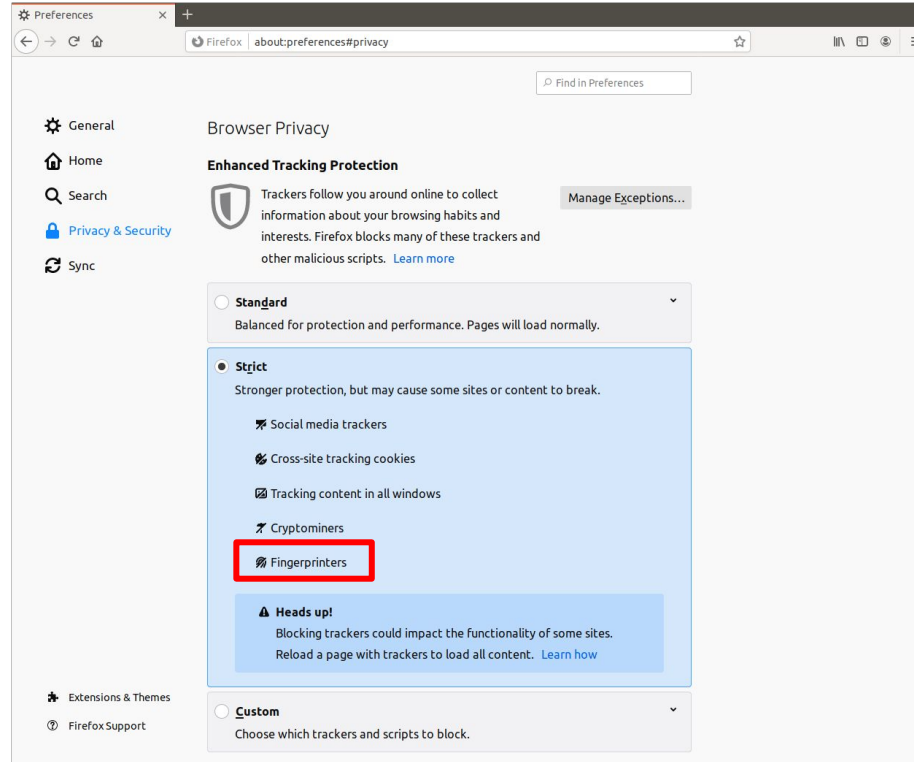
IP Address + Device
Properties



Identity

Fingerprinting blocking available since Firefox 67

Firefox 70



How can we prevent identity-based tracking?

Request: a safer way to do anti-fraud