CSE 484: Computer Security and Privacy

Web Security

[Web ~~Privacy~~]

tracking
ads
???

Winter 2021

David Kohlbrenner

dkohlbre@cs.washington.edu

Thanks to Franzi Roesner, Dan Boneh, Dieter Gollmann, Dan Halperin, Yoshi Kohno, Ada Lerner, John Manferdelli, John Mitchell, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials …

Admin

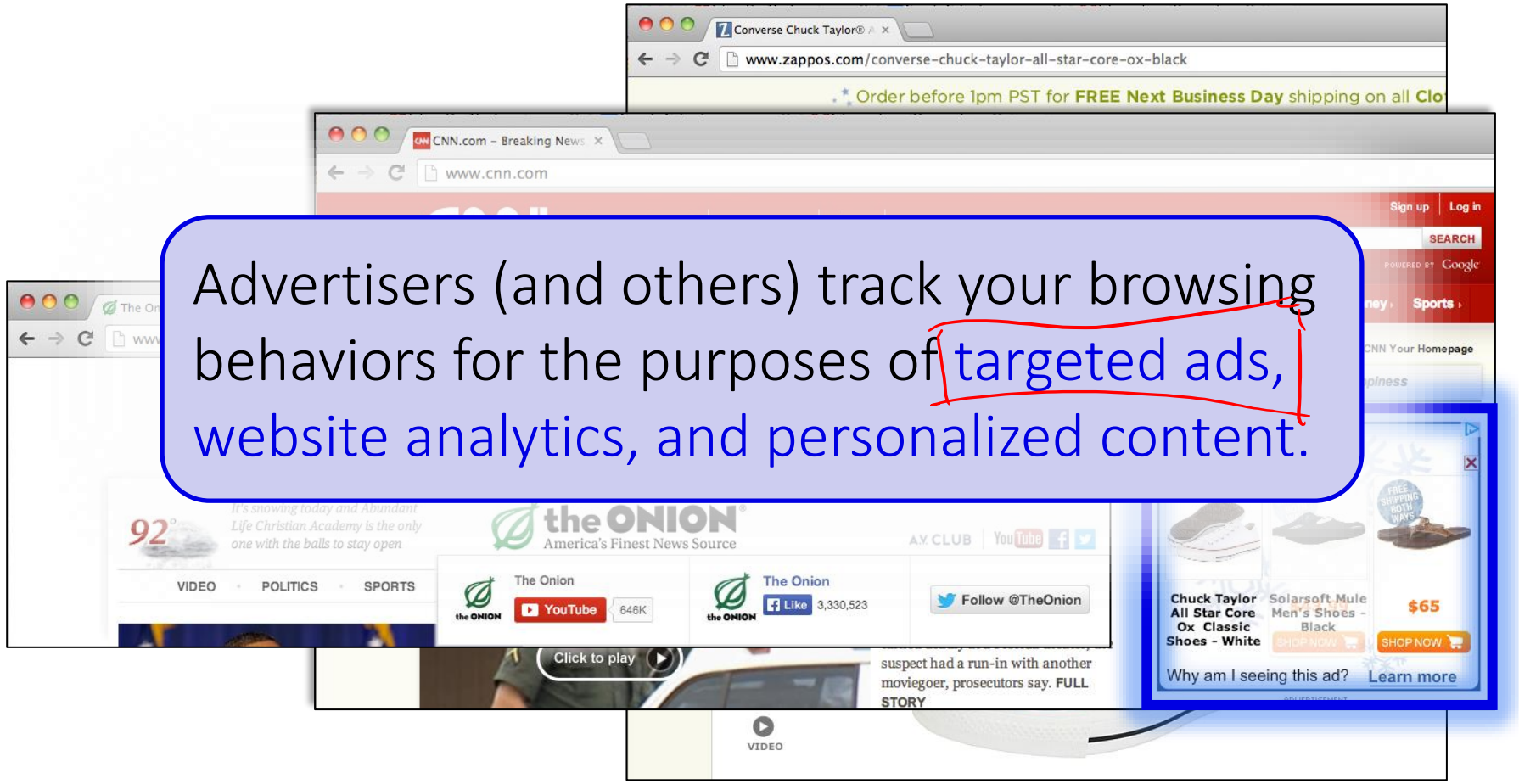
- Lab 2:
  - Check access ASAP
  - Read FAQs ☺
- Homework 3: out soon, due 03/5
- Guest lecture on Wednesday
  - Alex Gantman – Qualcomm – Embedded systems security and careers

not recorded

A topic in flux

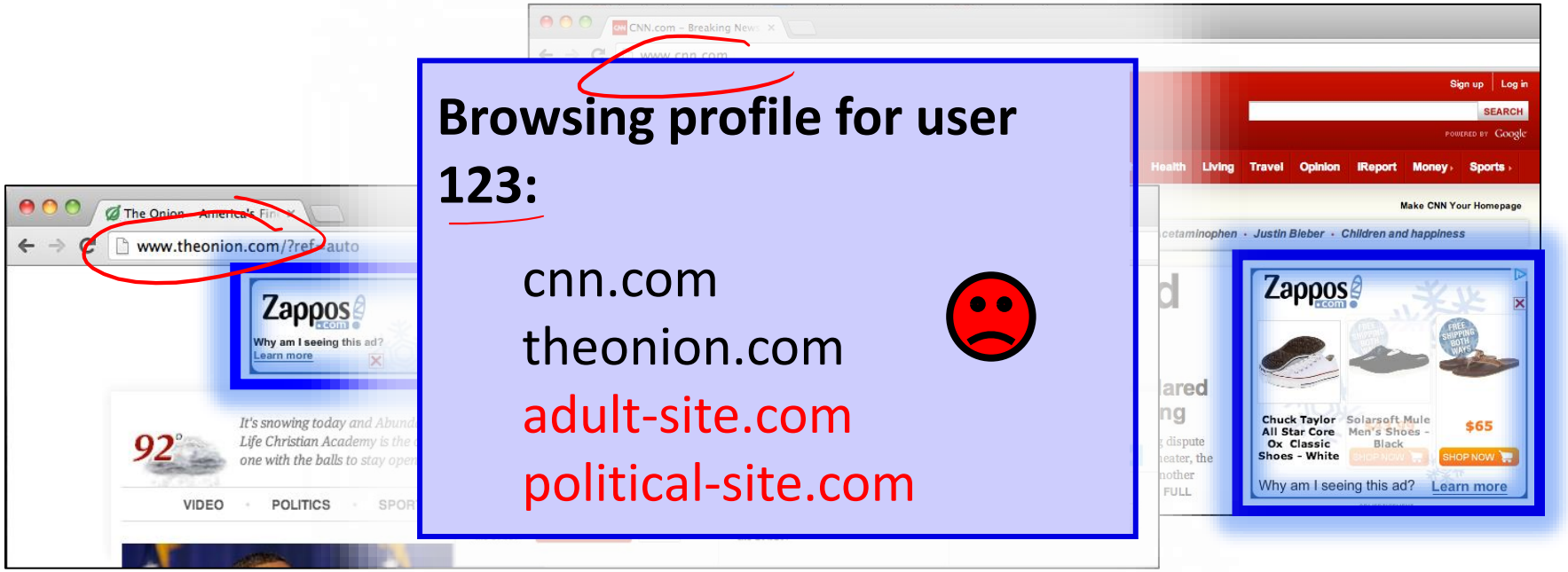
- Tracking via cookies
- Tracking via other methods
- Fingerprinting

Ads That Follow You



Advertisers (and others) track your browsing behaviors for the purposes of targeted ads, website analytics, and personalized content.

Third-Party Web Tracking



These ads allow **criteo.com** to link your visits between sites, even if you never click on the ads.

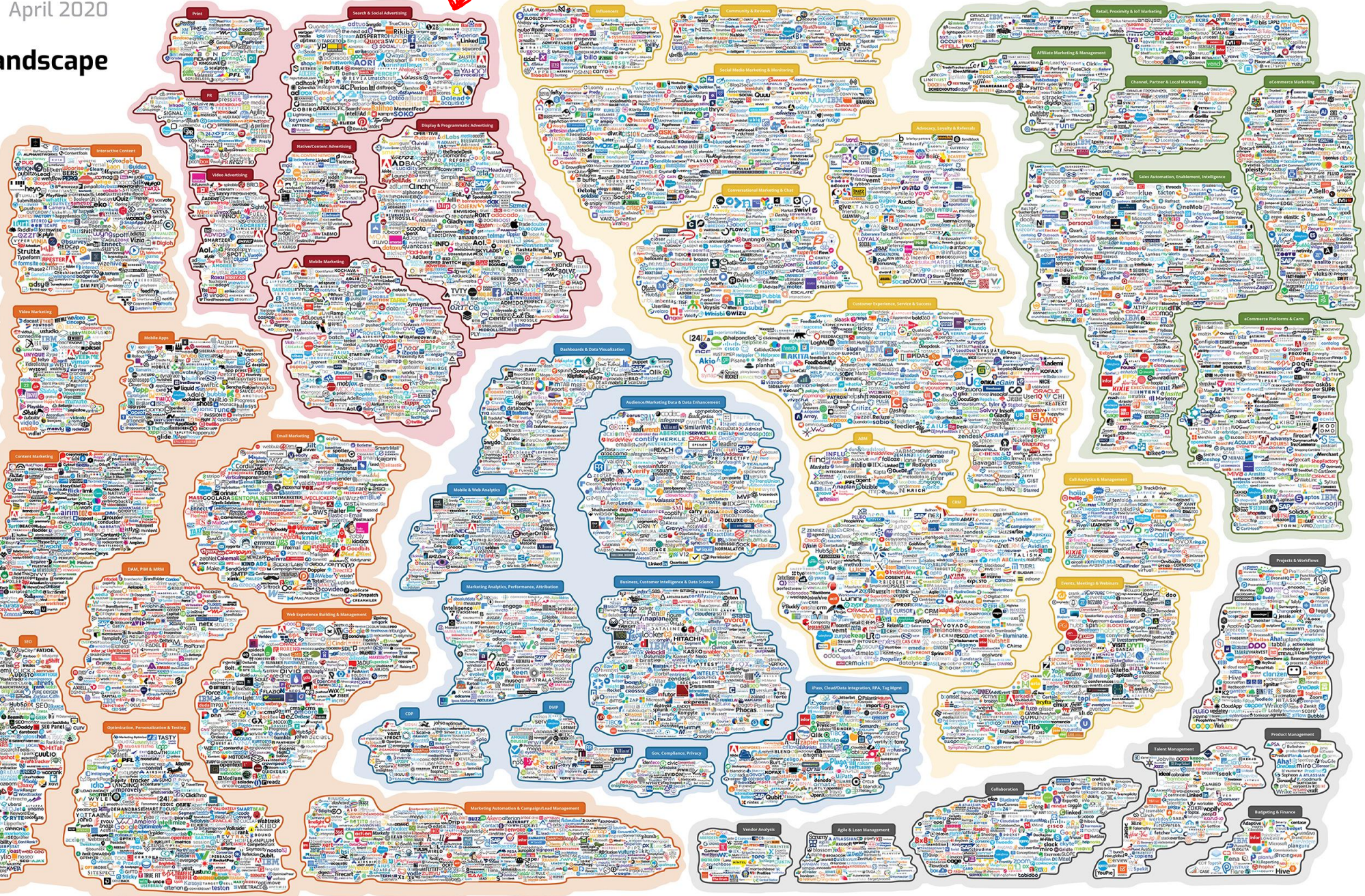
chiefmartec.com April 2020

Marketing Technology Landscape

The Martech 5000

Total Solutions	8,000
Advertising & Promotion	922
Content & Experience	1,936
Social & Relationships	1,969
Commerce & Sales	1,314
Data	1,258
Management	601

Access all the data of this landscape & more at **martech5000.com**



Produced by Scott Brinker (@chiefmartec) and Blue Green Brands (@bluegreenbrands).

Concerns About Privacy

THE WALL STREET JOURNAL.

WHAT THEY KNOW | JULY 30, 2010

The Web's New Gold Mine: Your Secrets

A Journal inv
bus

CNN

Your Privacy

Big
dep

By
Hid
all to be put up

The file consists
identifies her as

The New York Times

May 6, 2011, 5:01 pm | 3 Comments

'Do Not Track' Privacy Bill Appears in Congress

By TANZINA VEGA

And the privacy legislation just keeps on coming.

On Friday, two bills were introduced in Washington in support of a Do Not Track mechanism that would give users control over how much of their data was collected by advertisers and other online companies.

Log In

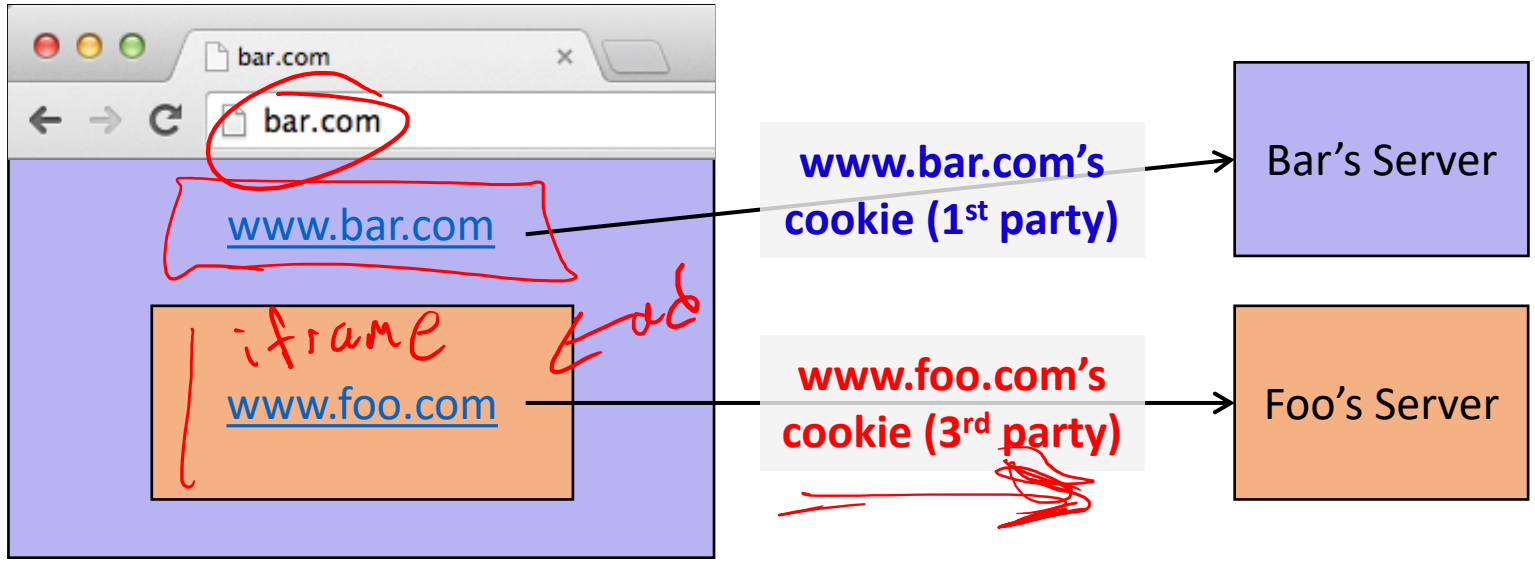
als
tion

By JENNIFER VALENTINO-DEVRIES,
JEREMY SINGER-VINE and ASHKAN SOLTANI
December 24, 2012

First and Third Parties

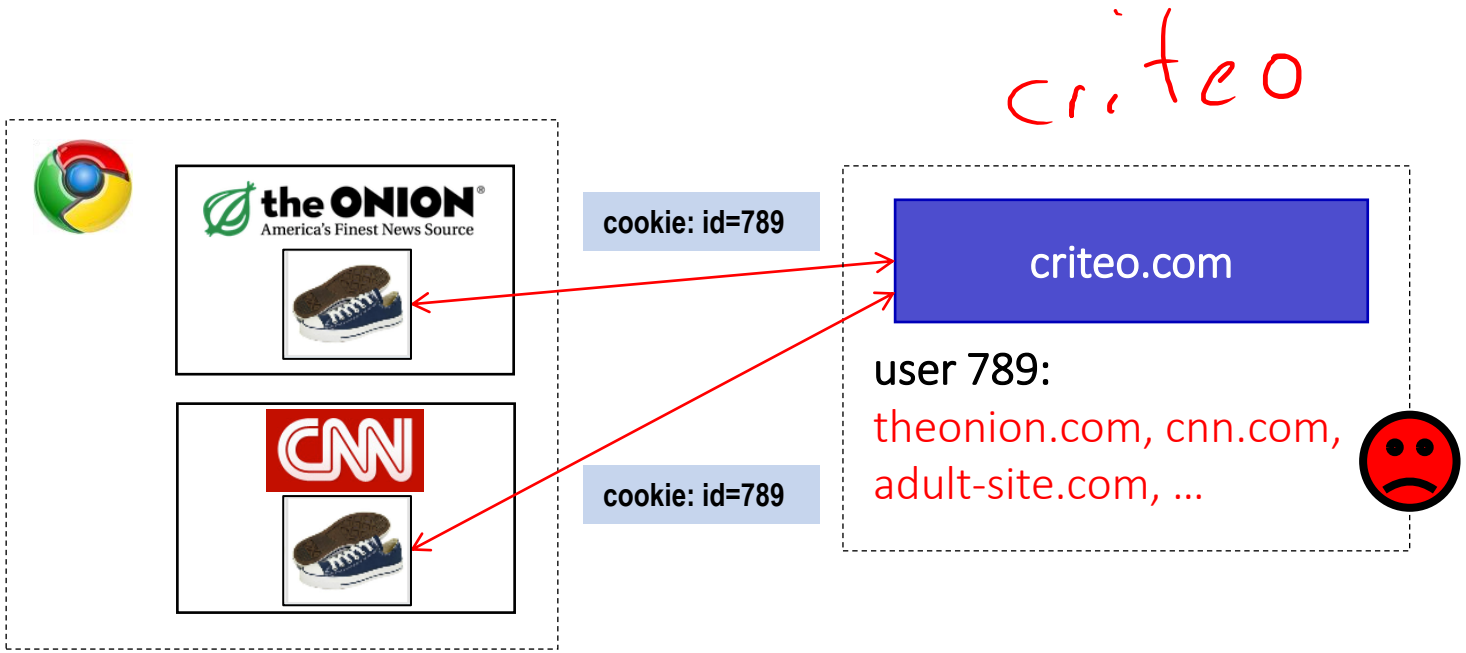
CSRFs

- First-party cookie: belongs to top-level domain.
- Third-party cookie: belongs to domain of embedded content (such as image, iframe).



Anonymous Tracking

Trackers included in other sites use third-party cookies containing unique identifiers to create browsing profiles.



Basic Tracking Mechanisms

- Tracking requires:
  - (1) re-identifying a user.
  - (2) communicating id + visited site back to tracker.

```
▽ Hypertext Transfer Protocol
  ▷ GET /pixel/p-3aud4J6uA4Z6Y.gif?labels=InvisibleBox&busty=2710 HTTP/1.1\r\n
    Host: pixel.quantserve.com\r\n
    Connection: keep-alive\r\n
    Accept: image/webp,*/*;q=0.8\r\n
    User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_2) AppleWebKit/537.36
    Referer: http://www.theonion.com/\r\n
    Accept-Encoding: gzip,deflate,sdch\r\n
    Accept-Language: en-US,en;q=0.8\r\n
    Cookie: mc=52a65386-f1de1-00ade-0b26e; d=ENkBRgGHD4GYEA35MMIL74MKiyDs1A2MQI1Q
```

GET

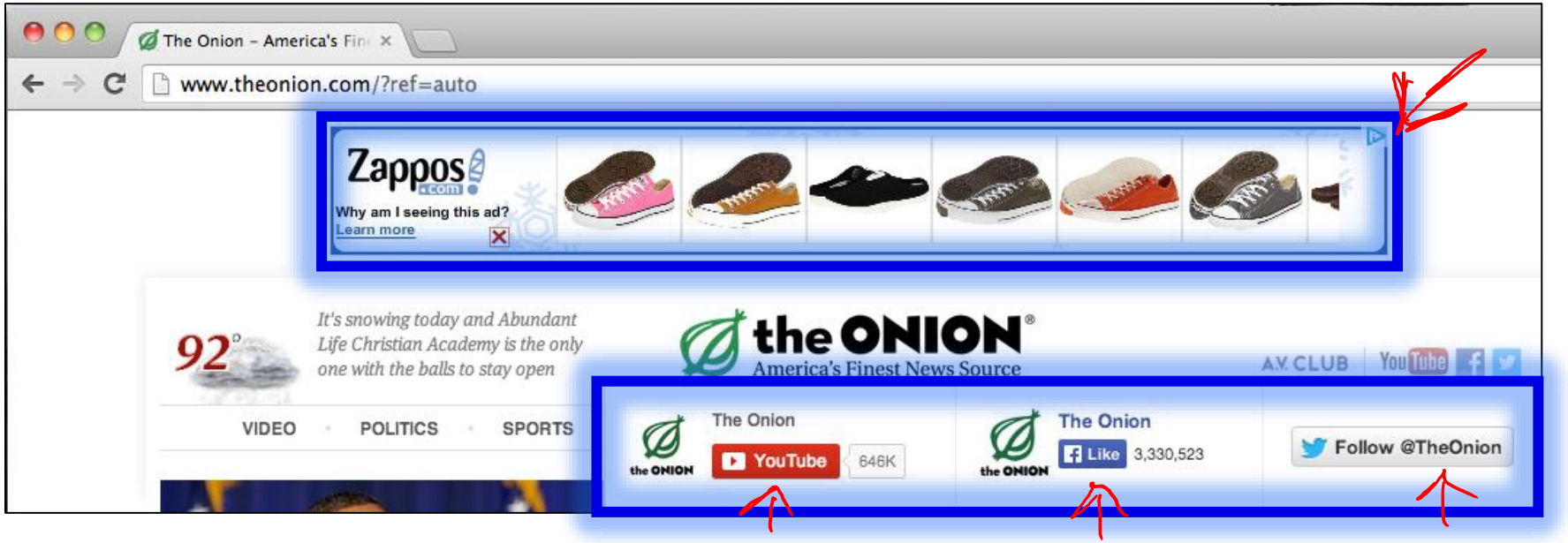
Tracking Technologies

- HTTP Cookies
- HTTP Auth
- HTTP Etags
- Content cache
- IE userData
- HTML5 protocol and content handlers
- HTML5 storage

- Flash cookies
- Silverlight storage
- TLS session ID & resume
- Browsing history
- window.name
- HTTP STS
- DNS cache

- "Zombie" cookies that respawn (http://samy.pl/evercookie)

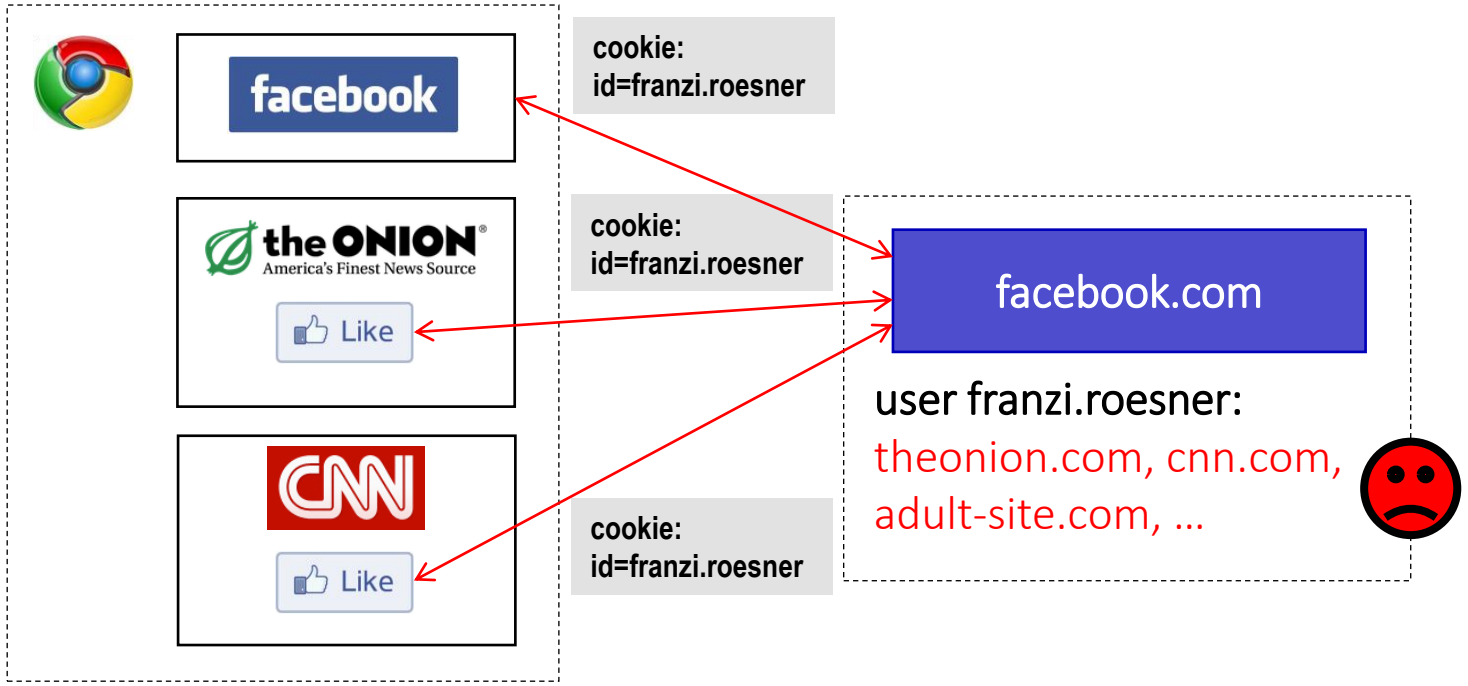
Other Trackers?



3rd party

"Personal" Trackers

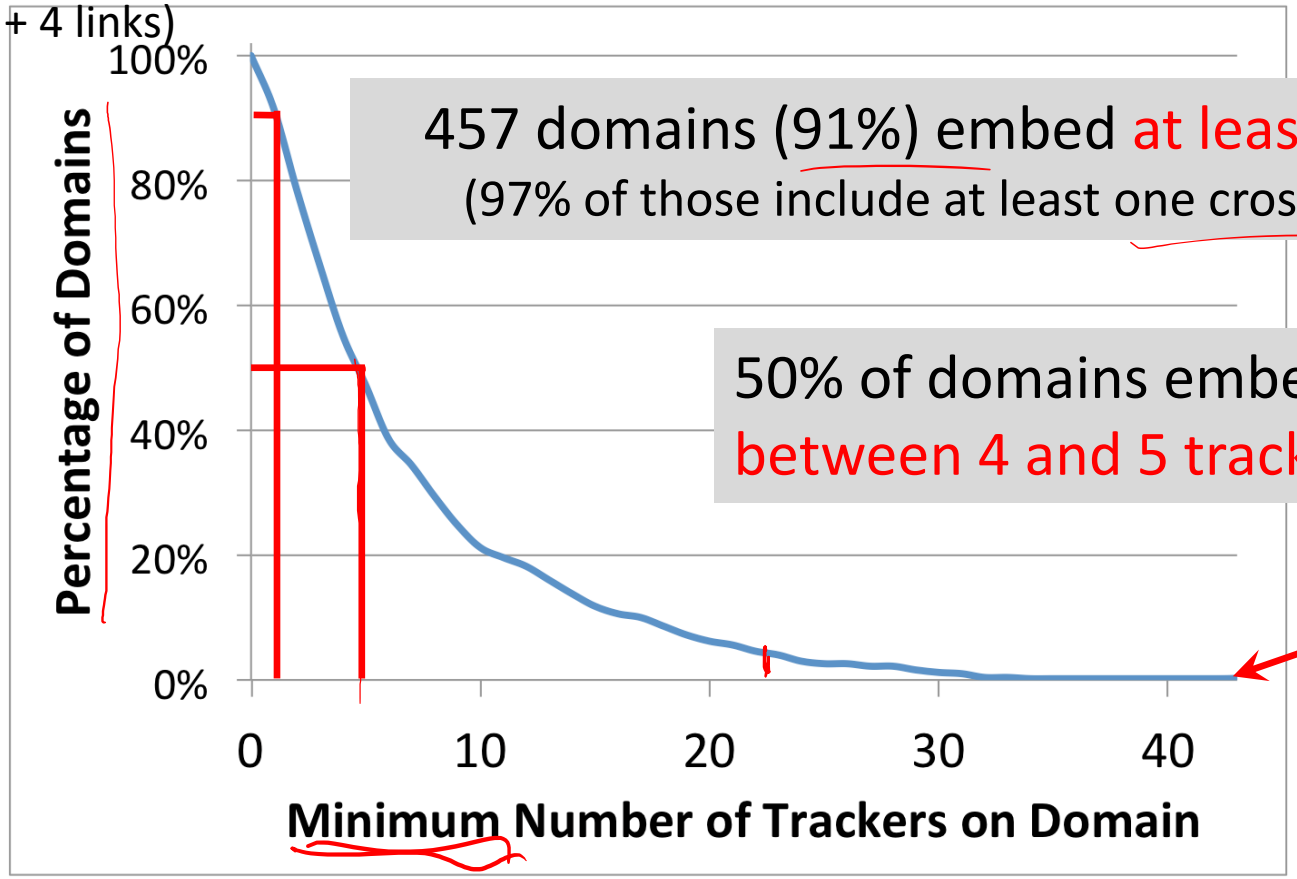
Personal Tracking



- Tracking is not anonymous (linked to accounts).
- Users directly visit tracker's site → evades some defenses.

How prevalent is tracking? (2011)

524 unique trackers on Alexa top 500 websites (homepages + 4 links)

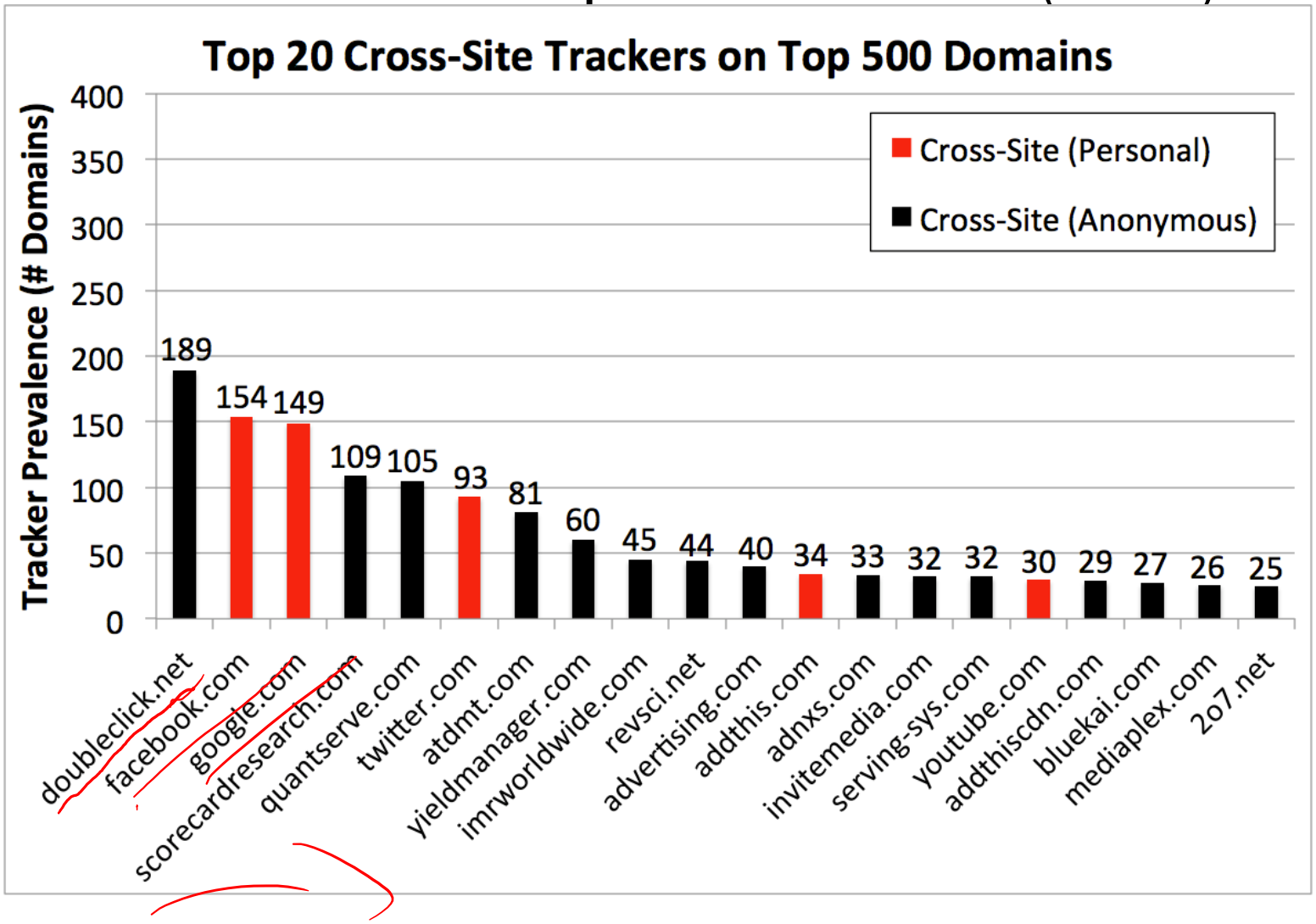


457 domains (91%) embed at least one tracker.
(97% of those include at least one cross-site tracker.)

50% of domains embed between 4 and 5 trackers.

One domain includes 43 trackers.

Who/what are the top trackers? (2011)

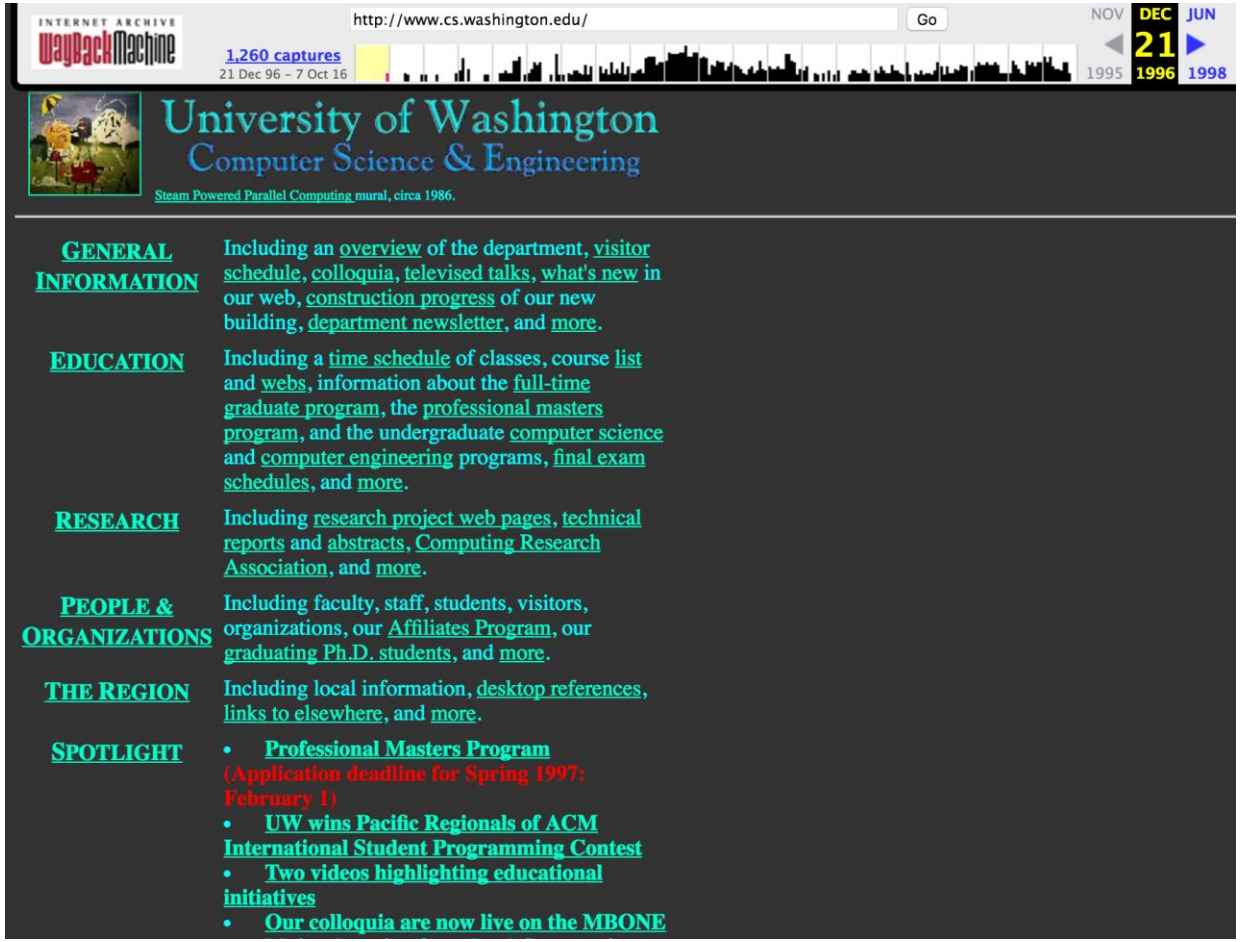


How has this changed over time?

- The web has existed for a while now…
  - What about tracking before 2011?
  - What about tracking before 2009?

- Solution: time travel!

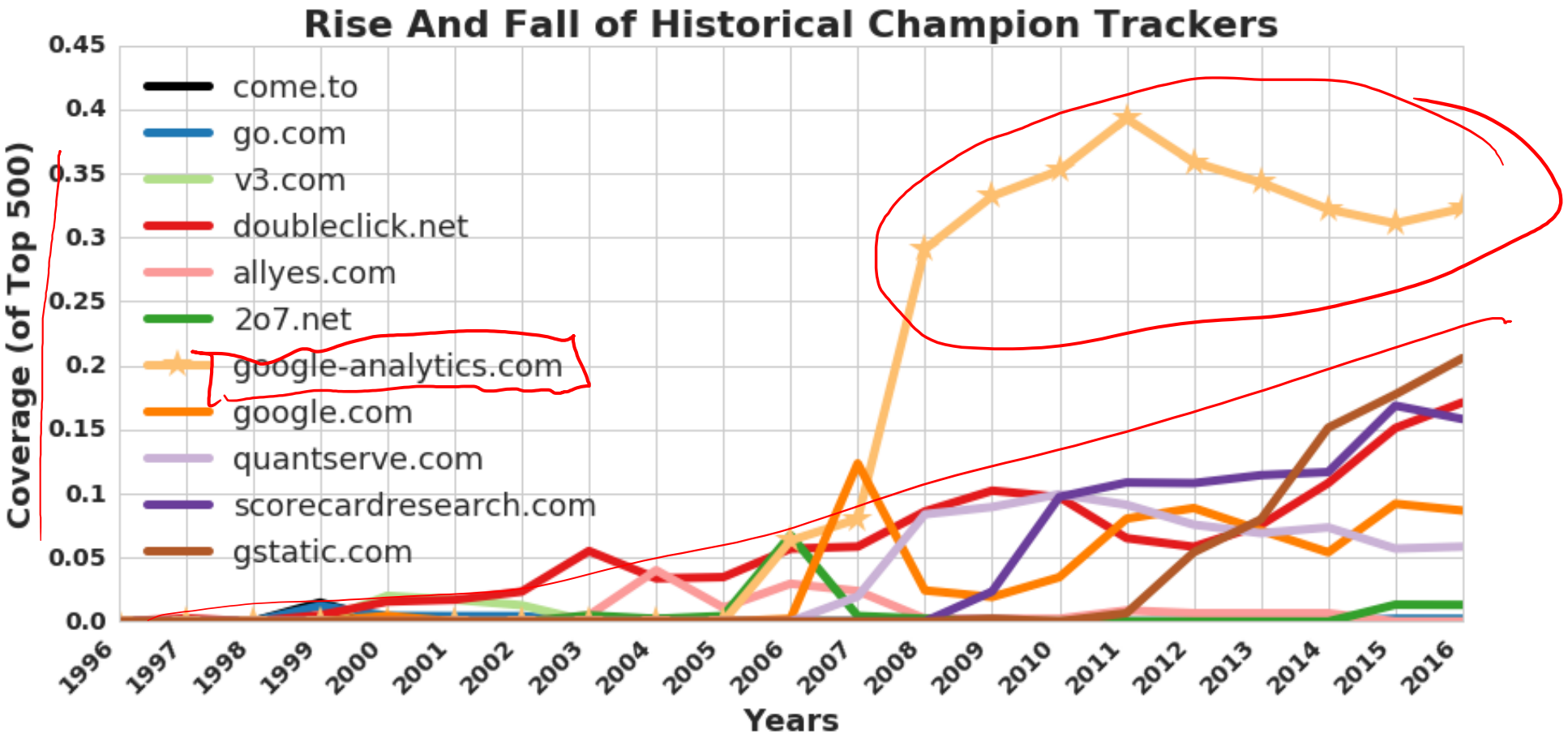
The Wayback Machine to the Rescue



Time travel for web tracking: http://trackingexcavator.cs.washington.edu

1996-2016: More & More Tracking

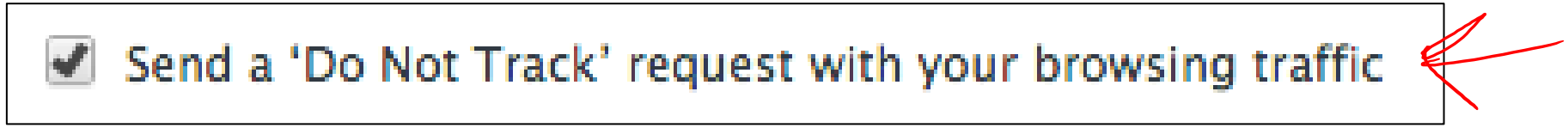
- More trackers of more types, more per site, more coverage



Defenses to Reduce Tracking

do not call

- Do Not Track?

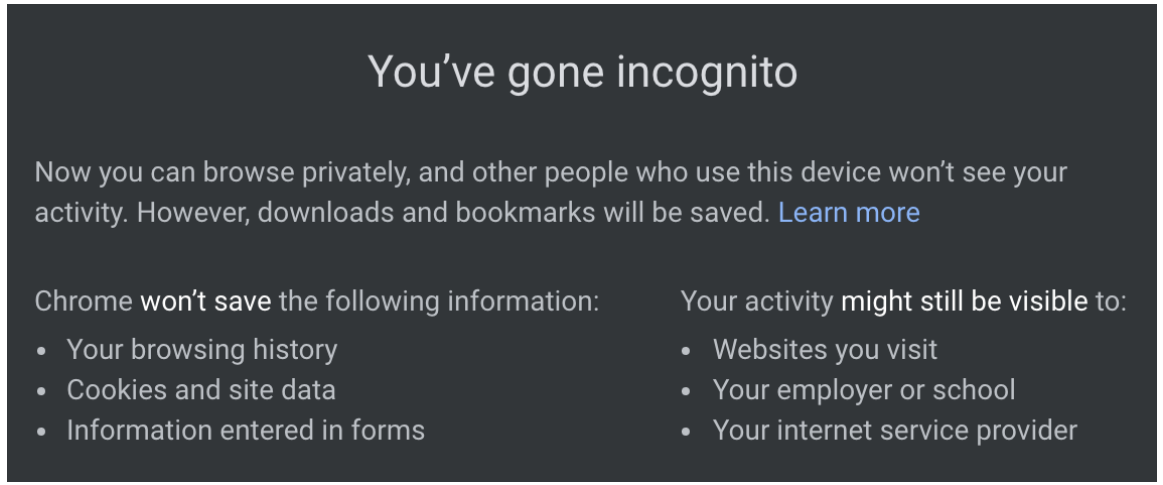


Do Not Track is not a technical defense: trackers must honor the request.

Defenses to Reduce Tracking

- Do Not Track proposal?
- Private browsing mode?

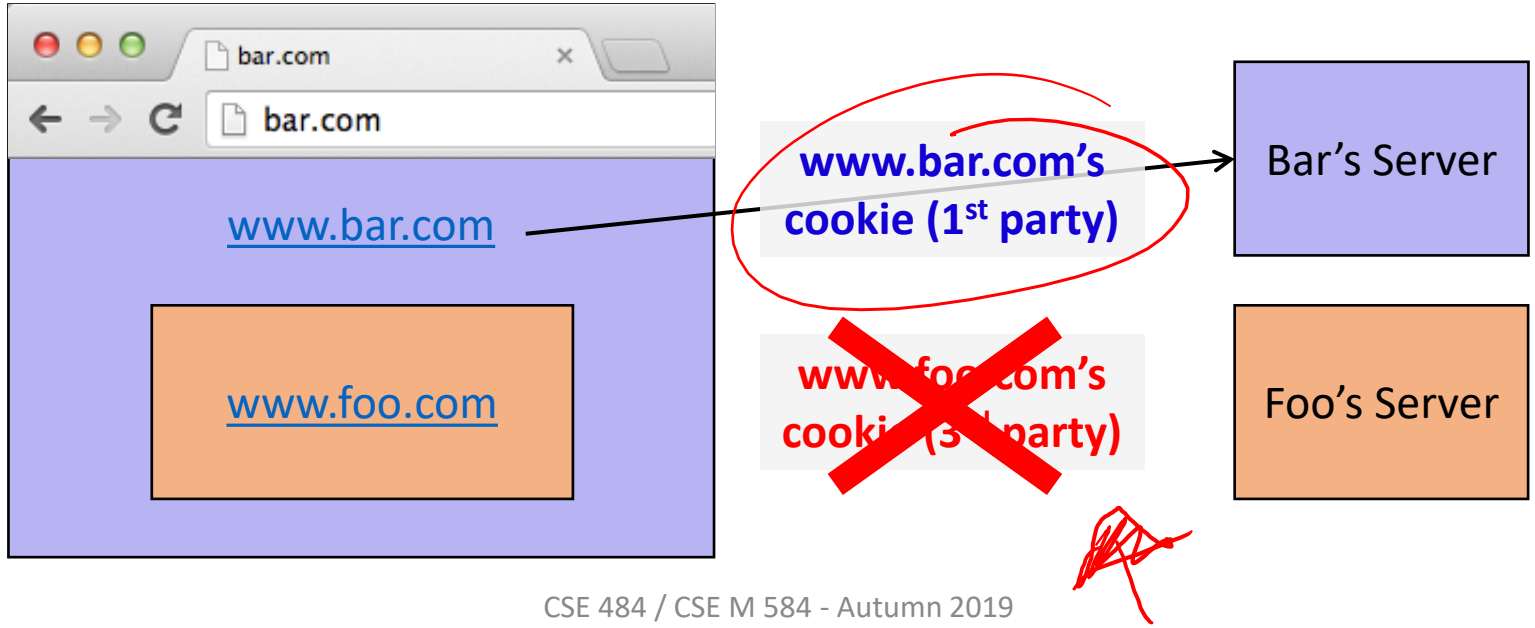
Private browsing mode protects against local, not network, attackers.



Defenses to Reduce Tracking

for browser

- Do Not Track proposal?
- Private browsing mode?
- Third-party cookie blocking?



Its real!

mid 2020

- Safari and FF (mostly) now block 3rd party cookies
  - https://webkit.org/blog/10218/full-third-party-cookie-blocking-and-more/
  - https://blog.mozilla.org/blog/2019/09/03/todays-firefox-blocks-third-party-tracking-cookies-and-cryptomining-by-default/
- Chrome… 2022?

> "By undermining the business model of many ad-supported websites, blunt approaches to cookies encourage the use of opaque techniques such as fingerprinting (an invasive workaround to replace cookies), which can actually reduce user privacy and control. We believe that we as a community can, and must, do better."

Fingerprinting is out there

- Better than a ‘voluntary’ cookie: involuntary, unchangeable id!
  - “Fingerprint”
- Idea: Measure ‘behavior’ of browser
  - Smash into unique ID

Fingerprinting Web Browsers

- User agent
- HTTP ACCEPT headers
- Browser plug-ins
- MIME support
- Clock skew
- Installed fonts
- Cookies enabled?
- Browser add-ons
- Screen resolution
- HTML5 canvas (differences in graphics SW/HW!)

HTML5 Canvas Fingerprinting - Text

Canvas

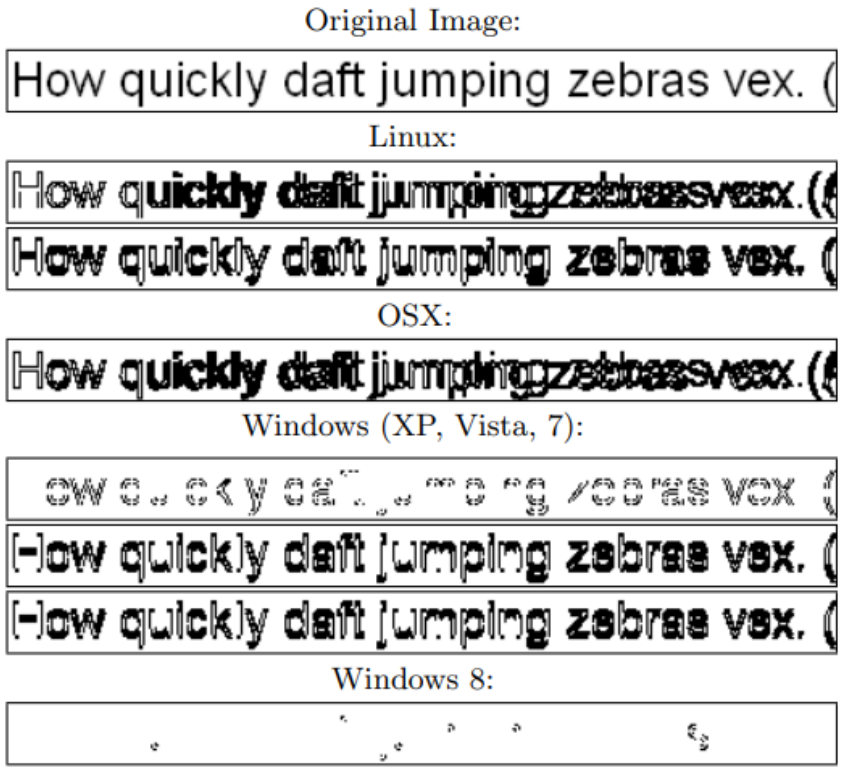
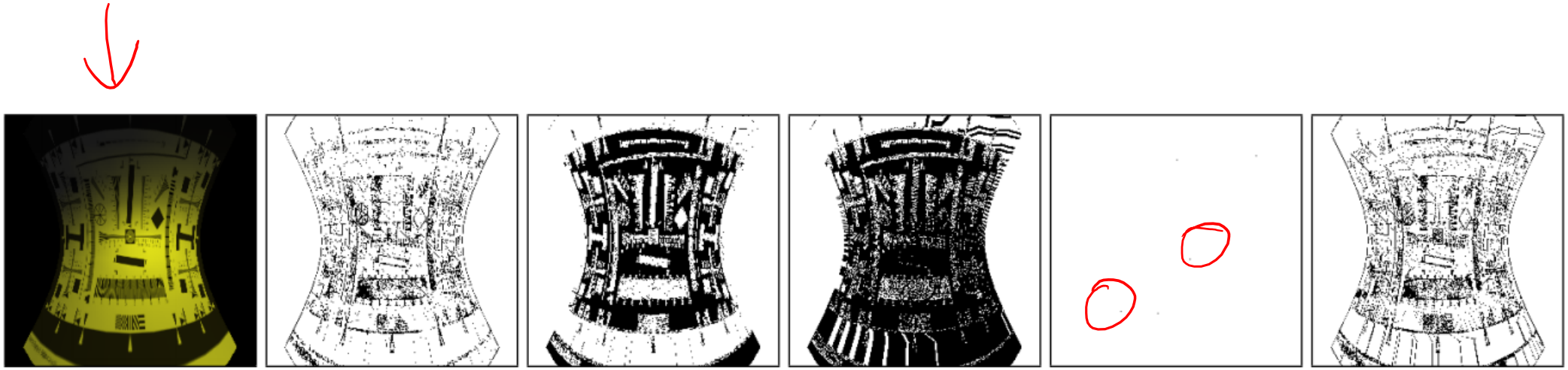


Figure 7: Difference maps for a group on `text_arial`

Mowery and Shacham, 2012

HTML5 Canvas Fingerprinting - Image



(a) Original (Intel G41) (b) Group 1 (Radeon HD 2400) (c) Group 20 (Intel 82945G) (d) Group 23 (Intel G33/G31) (e) Group 25 (Intel HD Graphics) (f) Group 36 (GeForce 6200)

Figure 10: Original render and difference maps for Group 24

Mowery and Shacham, 2012

And its out there!

2014

5%

Figure 4: Different images printed to canvas by fingerprinting scripts. Note that the phrase *"Cwm fjordbank glyphs vext quiz"* in the top image is a *perfect pangram*, that is, it contains all the letters of the English alphabet only once to maximize diversity of the outcomes with the shortest possible string.

Fingerprinting Web Browsers

- User agent
- HTTP ACCEPT headers
- Browser plug-ins
- MIME support
- Clock skew
- Installed fonts
- Cookies enabled?
- Browser add-ons
- Screen resolution
- HTML5 canvas (differences in graphics SW/HW!)

COVER YOUR TRACKS

eff

panopticlick

See how trackers view your browser

Learn About

HOW TO READ YOUR REPORT

You will see a summary of your overall tracking protection. The first section gives you a general idea of what your browser configuration is blocking (or not blocking). Below that is a list of specific browser characteristics in the format that a tracker would view them. We also provide descriptions of how they are incorporated into your fingerprint.

HOW CAN TRACKERS TRACK YOU?

Trackers use a variety of methods to identify and track users. Most often, this includes tracking cookies, but it can also include browser fingerprinting. Fingerprinting is a sneakier way to track users and makes it harder for users to regain control of their browsers. This report measures how easily trackers might be able to fingerprint your browser.

HOW CAN I USE MY RESULTS TO BE MORE ANONYMOUS?

Knowing how identifiable you are, or whether you are blocking trackers, can help you take steps to better protect your privacy. Browser add-ons or protection mechanisms built into the browser can help. Even so, the sneakiest trackers have ways around even the strongest security.

Here are your Cover Your Tracks results. They include an overview of how visible you are to trackers, with an index (and glossary) of all the metrics we measure below.

Our tests indicate that you have **strong protection against Web tracking**.

IS YOUR BROWSER:

Blocking tracking ads?	Yes
Blocking invisible trackers?	Yes
Protecting you from fingerprinting?	Your browser has a nearly-unique fingerprint

Still wondering how fingerprinting works?

LEARN MORE

Note: because ... easure all forms of tracking and p

Your R

One in 145,235 browsers have the same fingerprint

Within our dataset of several hundred thousand visitors tested in the past 45 days, only **one in 145235.5 browsers have the same fingerprint as yours.**

Fingerprinting as a security measure

- Blocking bots (e.g. reCAPTCHA)

3rd google cookies

- Validating users over-time

gmail

How should we view tracking and fingerprinting efforts?

“Privacy preserving” personalized ads

- https://github.com/WICG/turtledove
  - The browser, not the advertiser, holds the information about what the advertiser thinks a person is interested in.
  - Advertisers can serve ads based on an interest, but cannot combine that interest with other information about the person — in particular, with who they are or what page they are visiting.
  - Web sites the person visits, and the ad networks those sites use, cannot learn about their visitors' ad interests.

2021