

CSE 484 / CSE M 584: Web Privacy

Fall 2023

Franziska (Franzi) Roesner
franzi@cs

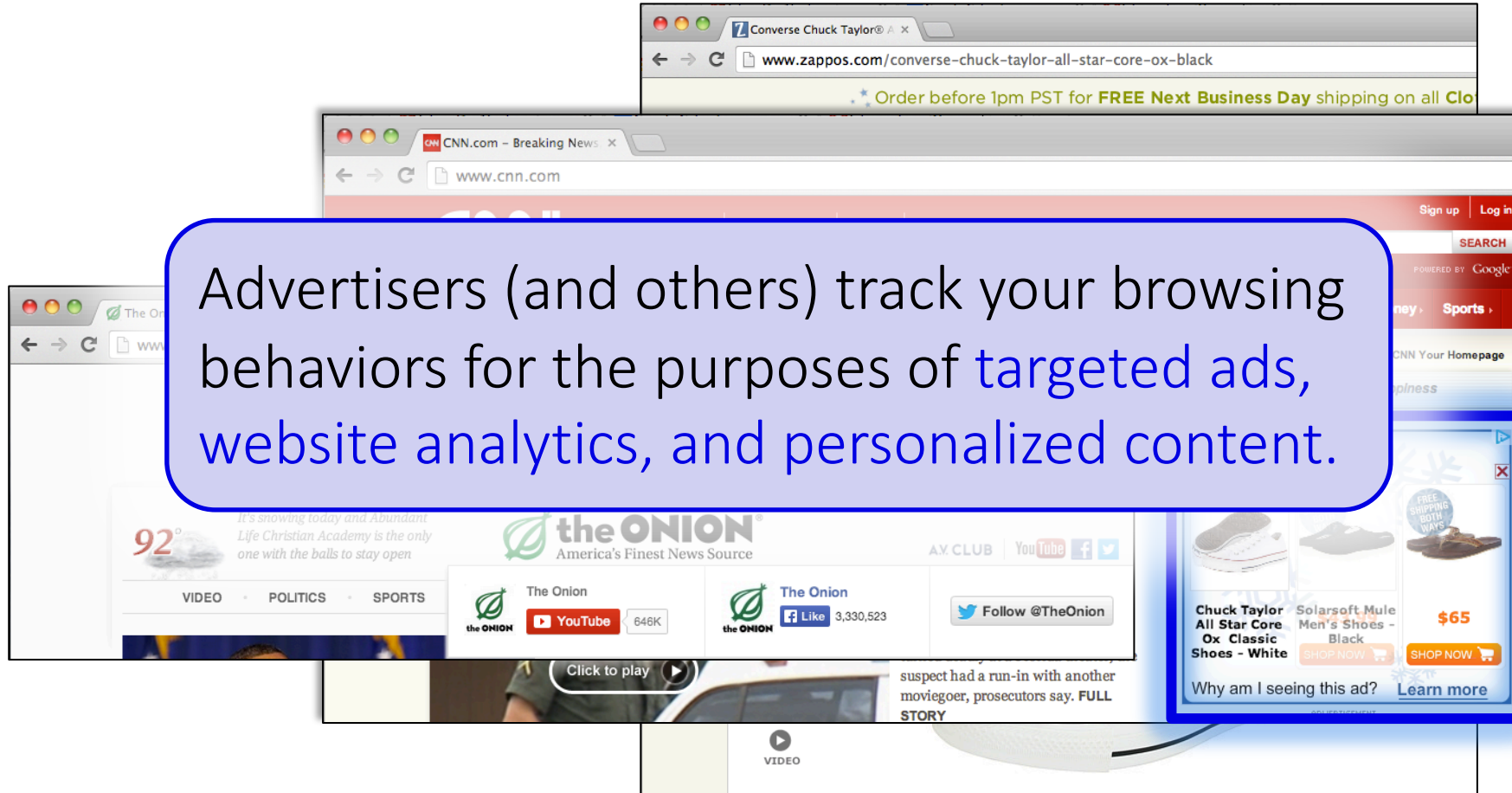
UW Instruction Team: David Kohlbrenner, Yoshi Kohno, Franziska Roesner. Thanks to Dan Boneh, Dieter Gollmann, Dan Halperin, John Manferdelli, John Mitchell, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials ...

Announcements

- Friday is a holiday, no class! (Veterans' Day)

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

Browsing profile for user 123:

- cnn.com
- theonion.com
- adult-site.com
- political-site.com

The image shows two browser windows. The left window is for 'The Onion - America's Finest' and the right is for 'CNN.com - Breaking News'. Both windows have a Zappos advertisement highlighted with a blue box. A central purple box contains a list of visited domains, with a red sad face icon to its right.

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

Concerns About Privacy

THE WALL STREET JOURNAL.

WHAT THEY KNOW | JULY 30, 2010

The Web's New Gold Mine: Your Secrets

A Journal
business c

Articl



Email

By Paolo Costa - 13 November 2018 431 0

Microtargeting and fake news: in the bubble of
online advertising

By JULIA ANGWIN

Hidden inside Ashley
all to be put up for sa

The file consists of a
identifies her as a 26-year-old female in Nashville, Tenn.

Your Privacy, For Sale

Big Data knows you're si
depressed

WHAT THEY KNOW

Websites Vary Prices, Deals
Based on Users' Information

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.

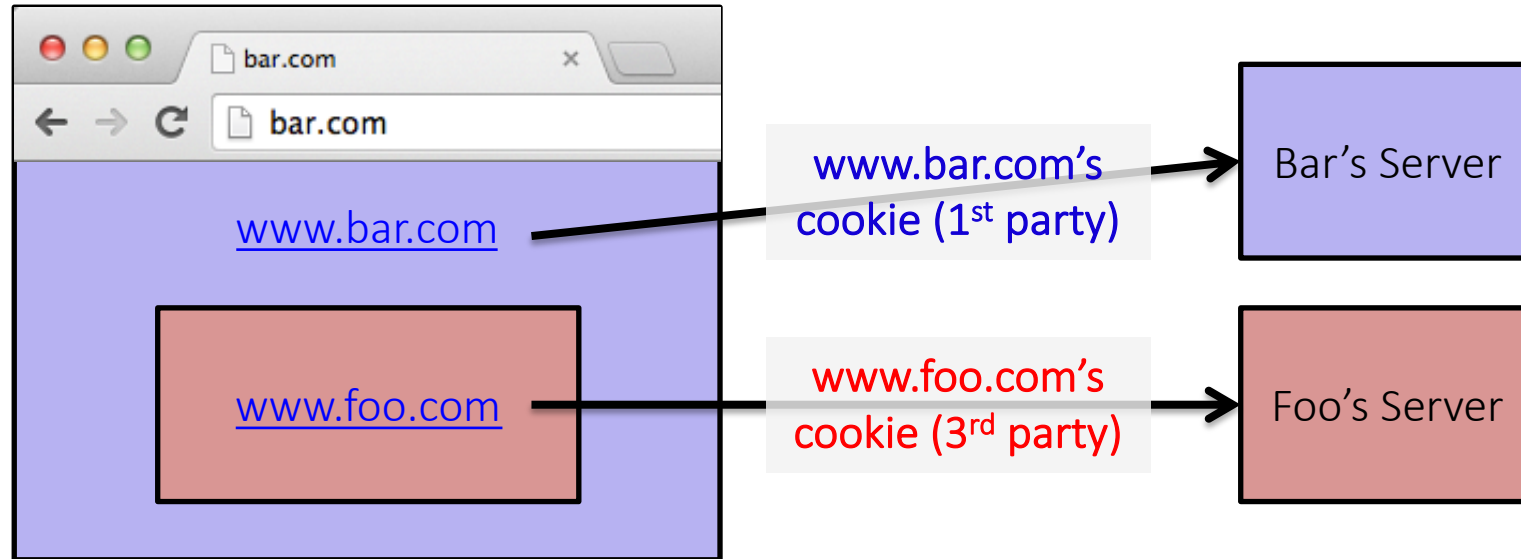
On Friday, two bills were introduced in Washington in support of a Do Not Track mechanism that would give users control over how much of their data was collected by advertisers and other online companies.

Outline

1. Understanding web tracking
2. Measuring web tracking
3. Defenses

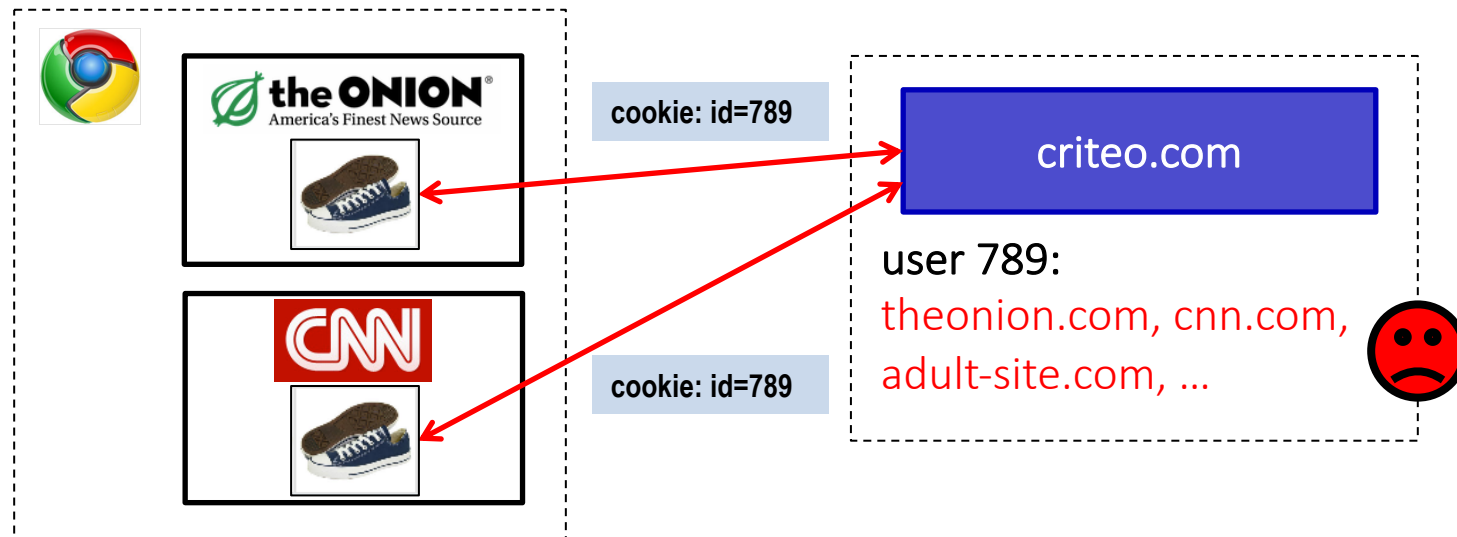
First and Third Parties

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



Cookies (etc.) Enable “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&bust=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\r\n
    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
```

Lots of Places to Store Identifiers...

- 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>)

You Don't Need Identifiers: Fingerprinting

- 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!)
- Etc. etc.

MY BROWSER FINGERPRINT

SEE YOUR BROWSER FINGERPRINT PROPERTIES

<https://www.amiunique.org/fingerprint>

ARE YOU UNIQUE ?

DOWNLOAD

TIMELINE

TODAY

7 DAYS

15 DAYS

30 DAYS








90 DAYS

ALL TIME

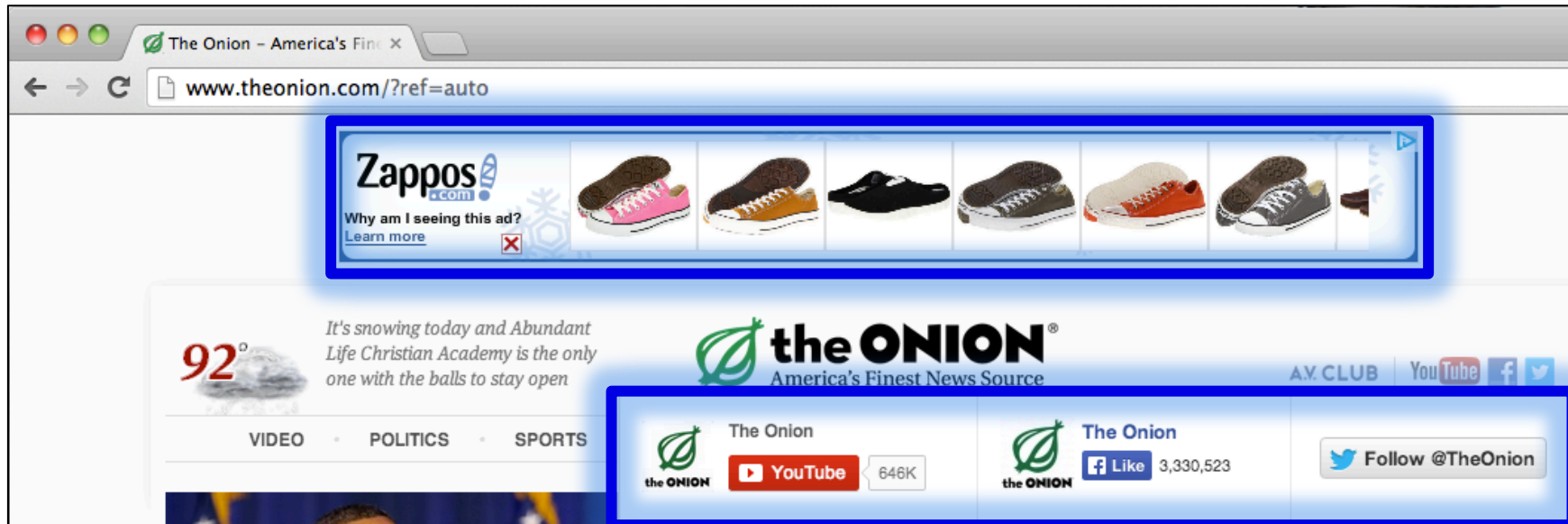
Yes! You are unique among the 2168672 fingerprints in our entire dataset.

JAVASCRIPT ATTRIBUTES

Search for an attribute

Attribute	Similarity ratio	Value
1 - User agent 	0.01 %	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/119.0.0.0 Safari/537.36
2 - Platform 	9.64 %	MacIntel
3 - Cookies enabled 	88.40 %	✓
4 - Timezone 	1.48 %	UTC-08:00
5 - Content language 	0.03 %	en-US,de-DE,de,en
6 - Canvas 	0.35 %	<small>Cwm fjordbank glyphs vext quiz, 😊</small> Cwm fjordbank glyphs vext quiz, 😊
7 - List of fonts (JS) 	0.00 %	Abadi MT Condensed Extra Bold Abadi MT Condensed Light Al Bayan Al Nile Al Tarikh And 287 others

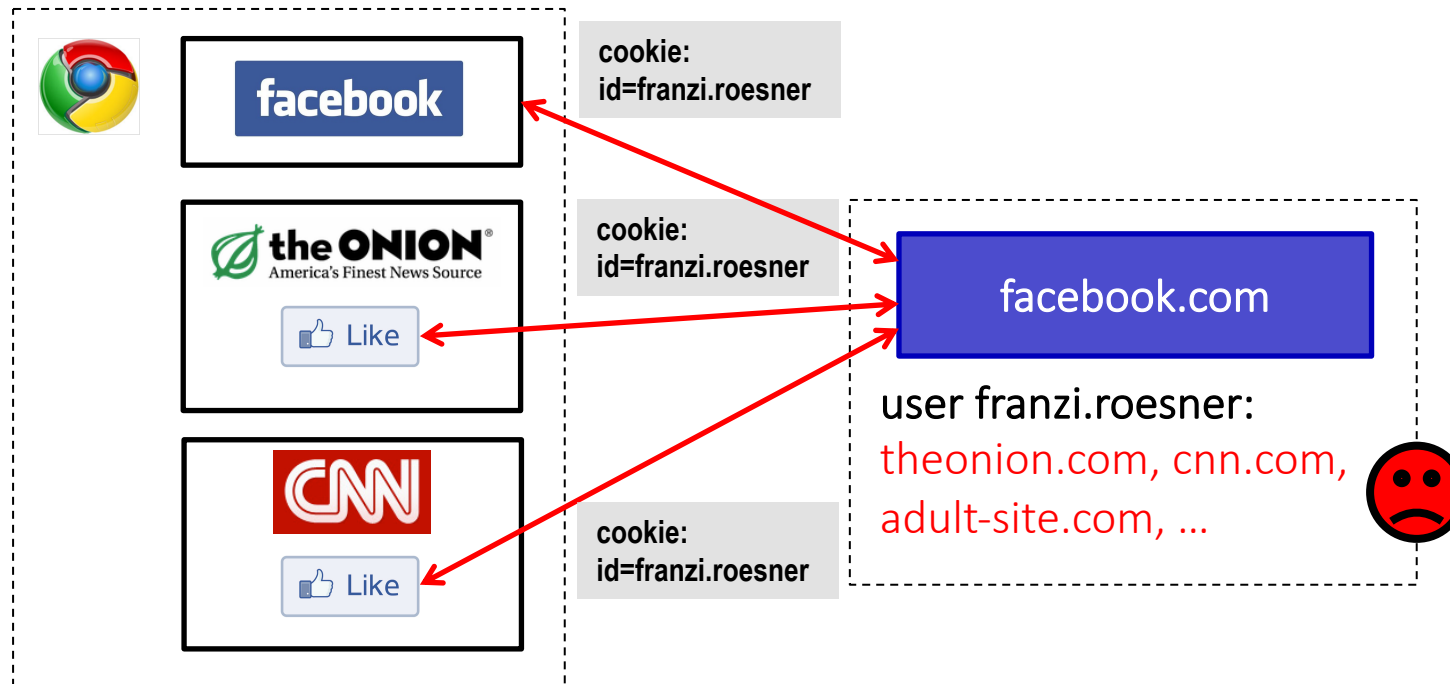
Other Trackers?



“Personal” Trackers



Personal Tracking



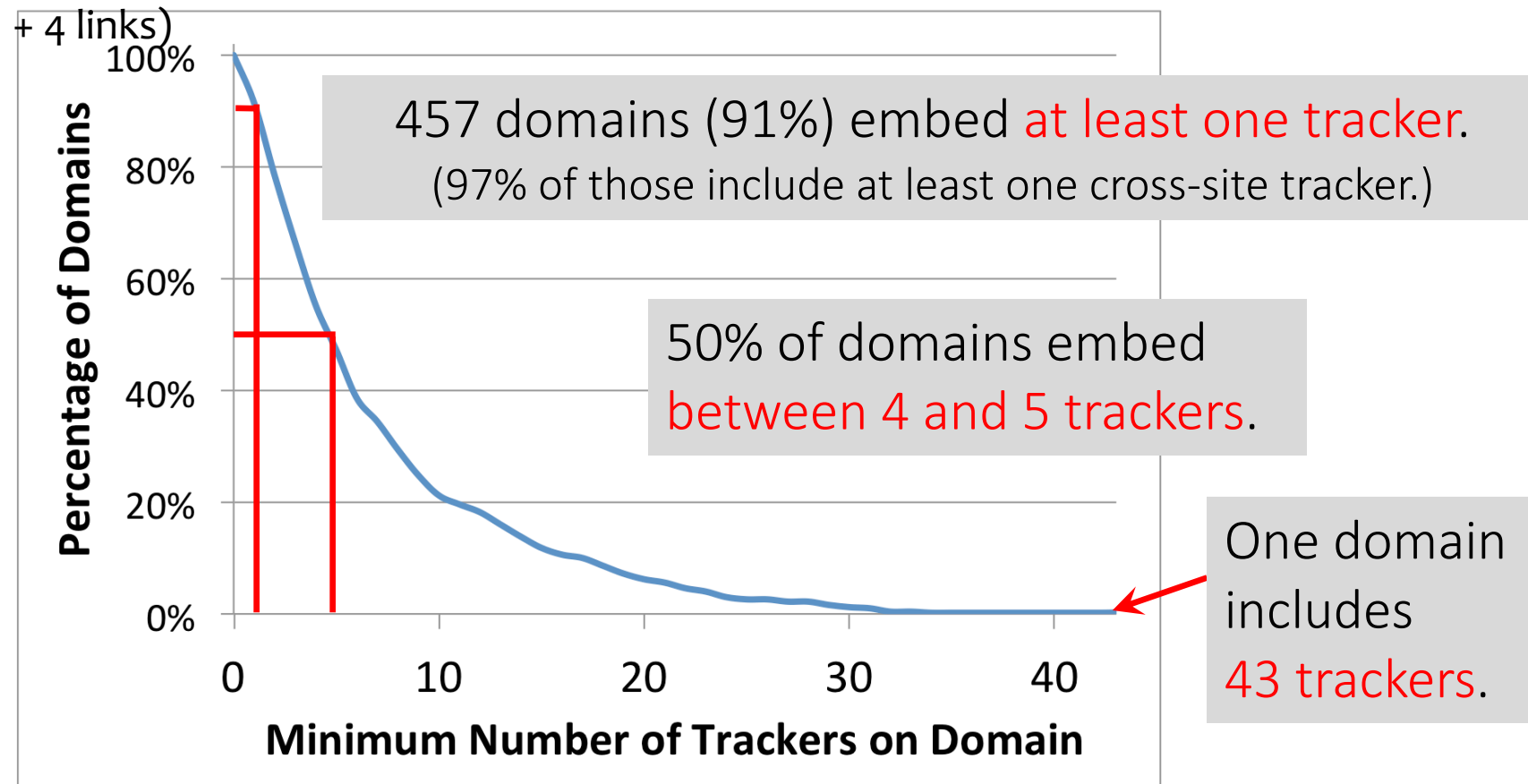
- Tracking is **not anonymous** (linked to accounts).
- Users **directly visit tracker's site** → evades some defenses.

Outline

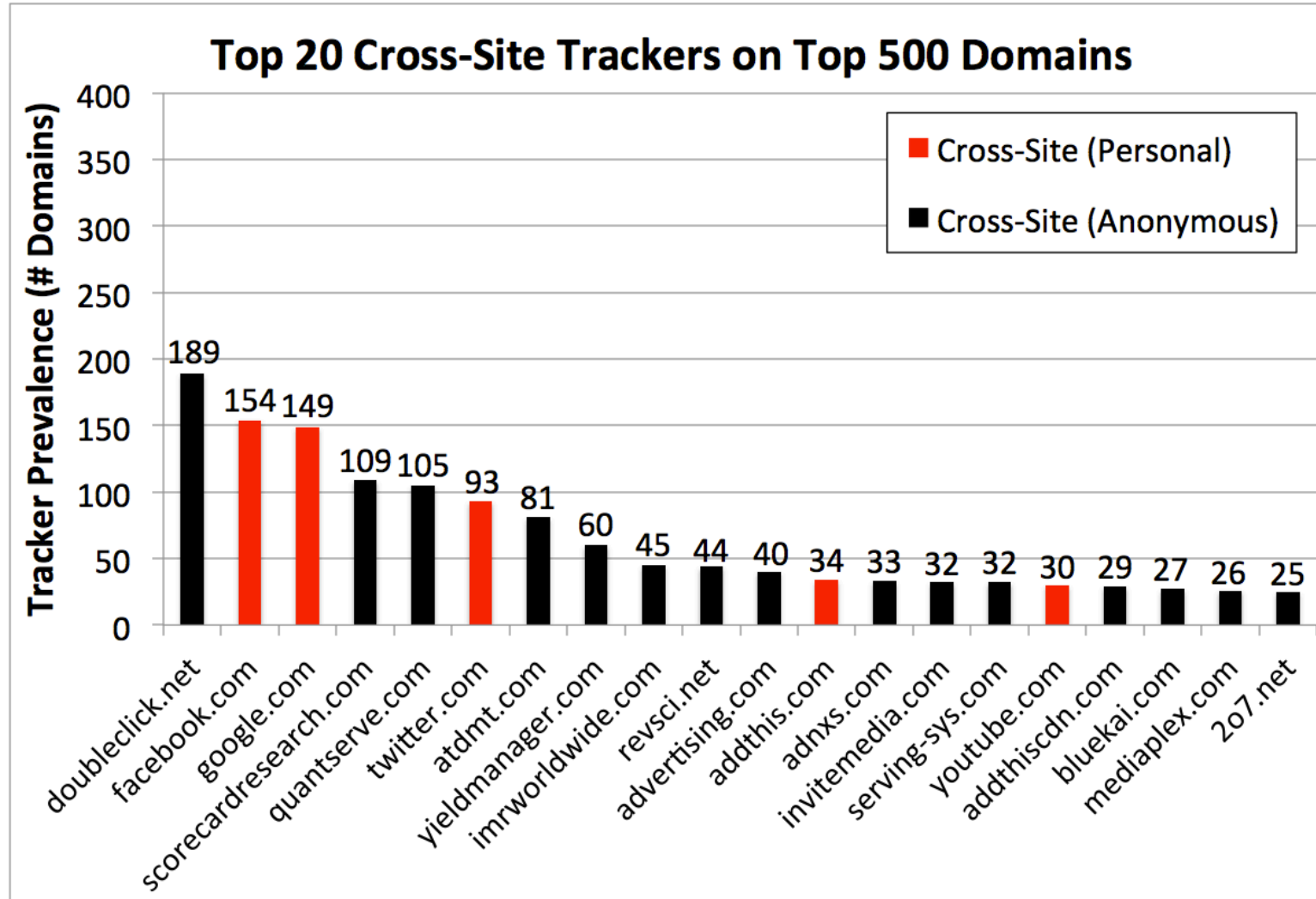
1. Understanding web tracking
2. Measuring web tracking
3. Defenses

How prevalent is tracking? (some years ago)

524 unique trackers on Alexa top 500 websites (homepages



Who/what are the top trackers? (some years ago)

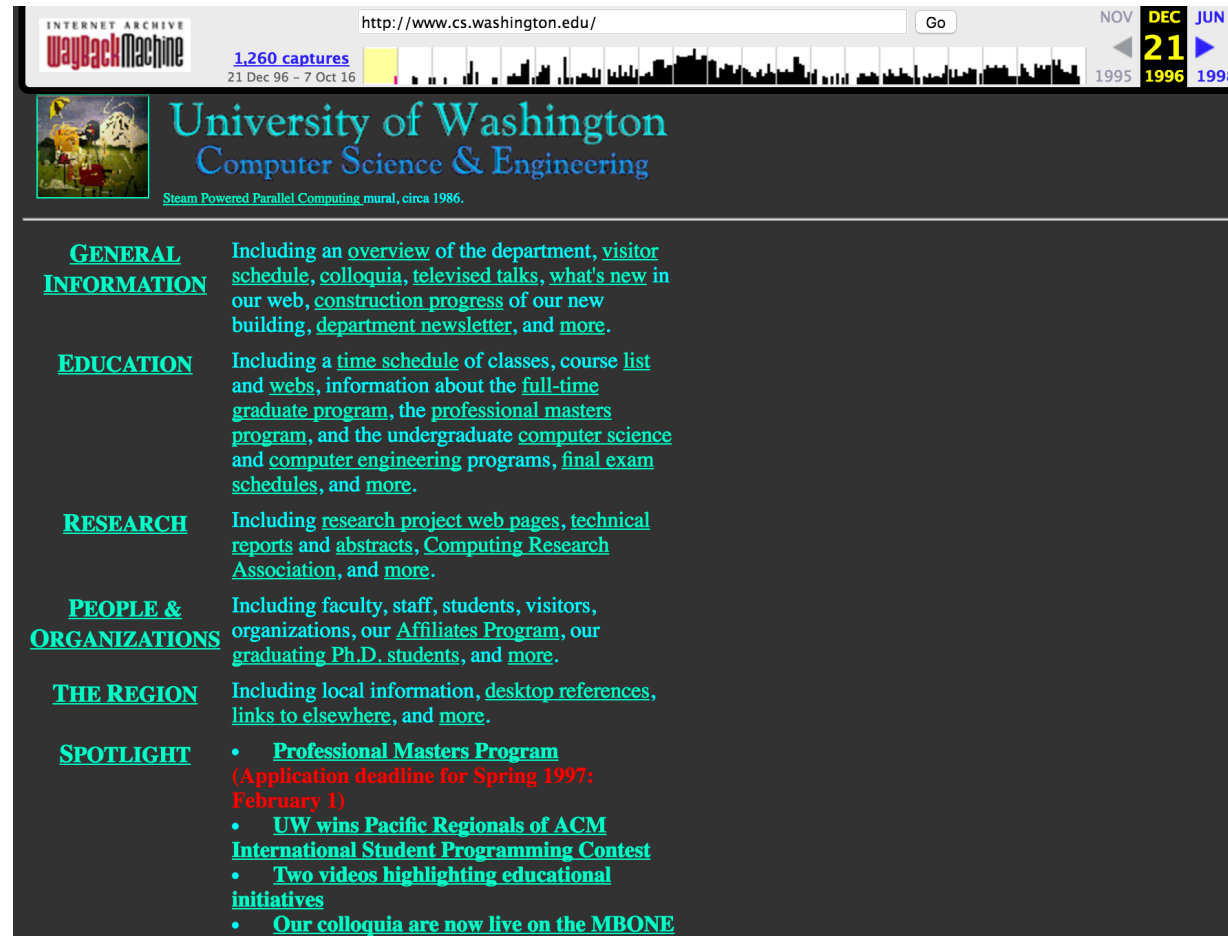


How has this changed over time?

- The web has existed for a while now...
 - What about tracking before 2011? (our first study)
 - What about tracking before 2009? (first academic study)
- Solution: **time travel!**



The Wayback Machine to the Rescue



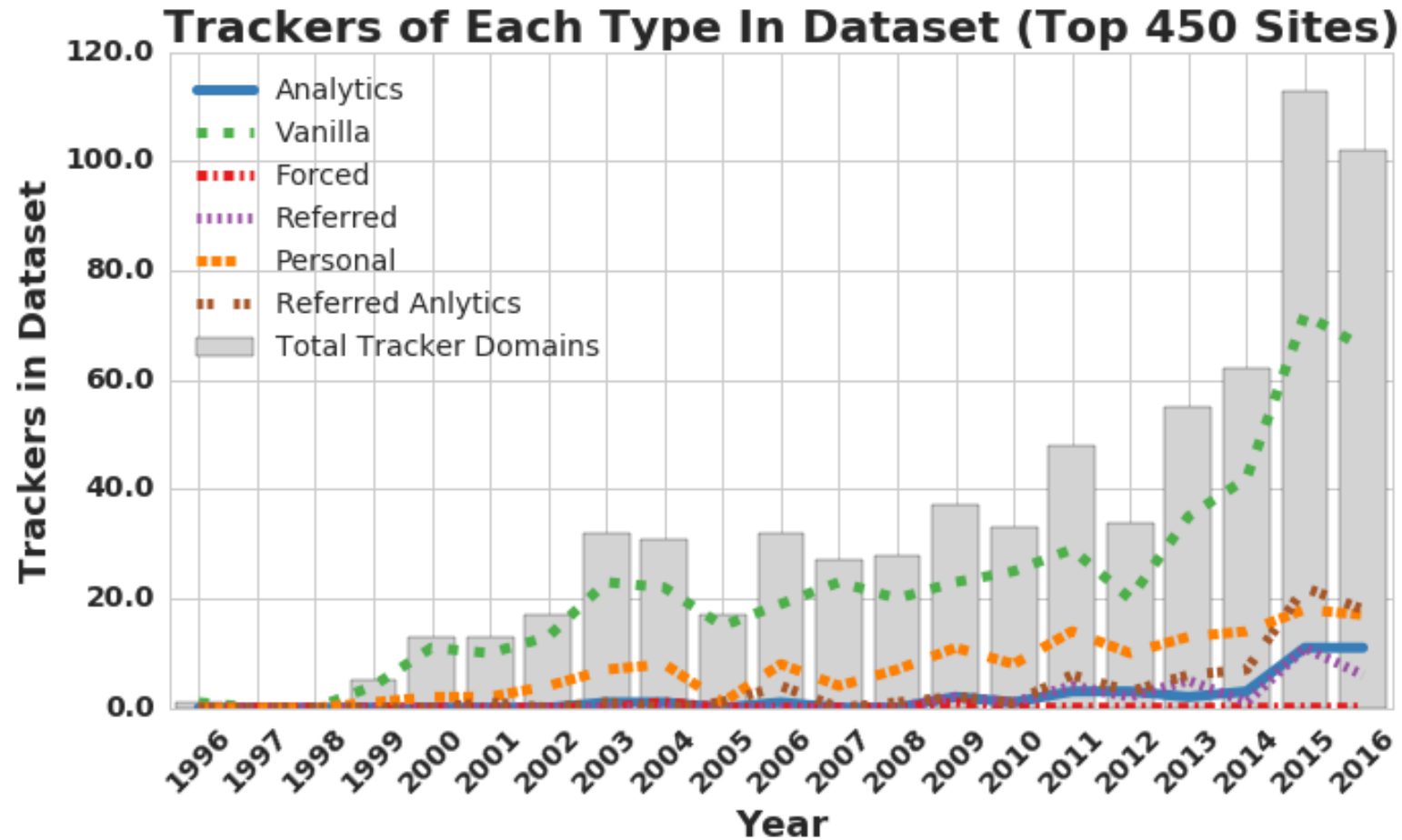
The screenshot shows the Wayback Machine interface for the URL <http://www.cs.washington.edu/>. The interface includes a search bar with the text "Go", a calendar showing the date "21" in "DEC" 1996, and a "1,260 captures" indicator. Below the header, there is a logo for the University of Washington Computer Science & Engineering and a small image of a mural titled "Steam Powered Parallel Computing mural, circa 1986". The main content area is divided into several sections:

- GENERAL INFORMATION**: Including an [overview](#) of the department, [visitor schedule](#), [colloquia](#), [televised talks](#), [what's new in our web](#), [construction progress](#) of our new building, [department newsletter](#), and [more](#).
- EDUCATION**: Including a [time schedule](#) of classes, [course list](#) and [webs](#), information about the [full-time graduate program](#), the [professional masters program](#), and the undergraduate [computer science](#) and [computer engineering](#) programs, [final exam schedules](#), and [more](#).
- RESEARCH**: Including [research project web pages](#), [technical reports](#) and [abstracts](#), [Computing Research Association](#), and [more](#).
- PEOPLE & ORGANIZATIONS**: Including faculty, staff, students, visitors, organizations, our [Affiliates Program](#), our [graduating Ph.D. students](#), and [more](#).
- THE REGION**: Including local information, [desktop references](#), [links to elsewhere](#), and [more](#).
- SPOTLIGHT**:
 - **Professional Masters Program** (Application deadline for Spring 1997: February 1)
 - **UW wins Pacific Regionals of ACM International Student Programming Contest**
 - **Two videos highlighting educational initiatives**
 - **Our colloquia are now live on the MBONE**

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

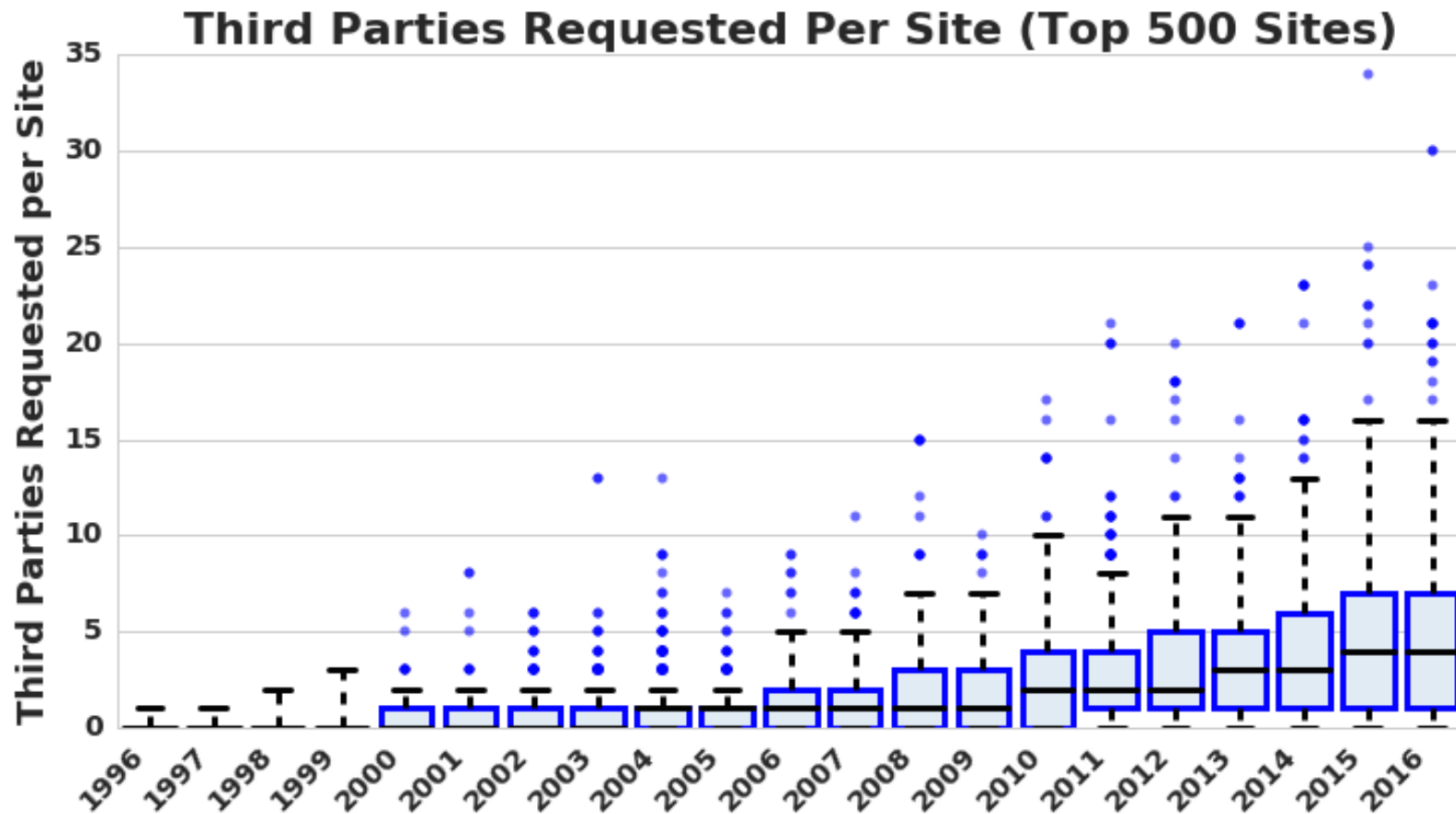
1996-2016: More & More Tracking

- More trackers of more types



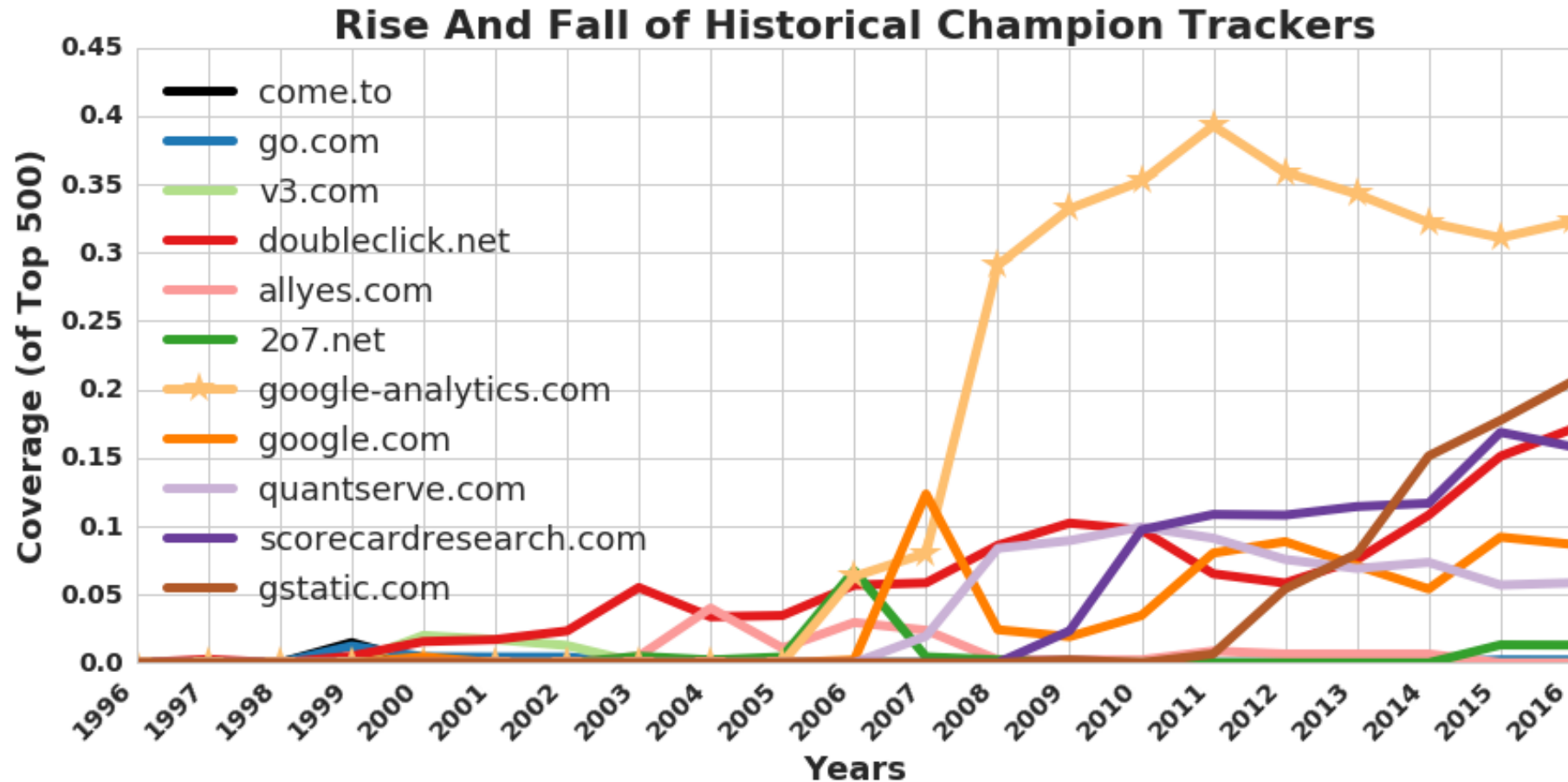
1996-2016: More & More Tracking

- More trackers of more types, [more per site](#)



1996-2016: More & More Tracking

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



Outline

1. Understanding web tracking
2. Measuring web tracking
3. Defenses

Defenses to Reduce Tracking

- Do Not Track proposal?

Send a 'Do Not Track' request with your browsing traffic

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

Defenses to Reduce Tracking

- Do Not Track proposal?
- Private browsing mode?

Mostly protects against local, not network, attackers.

You've gone incognito

Now you can browse privately, and other people who use this device won't see your activity. However, downloads and bookmarks will be saved. [Learn more](#)

Chrome won't save the following information:

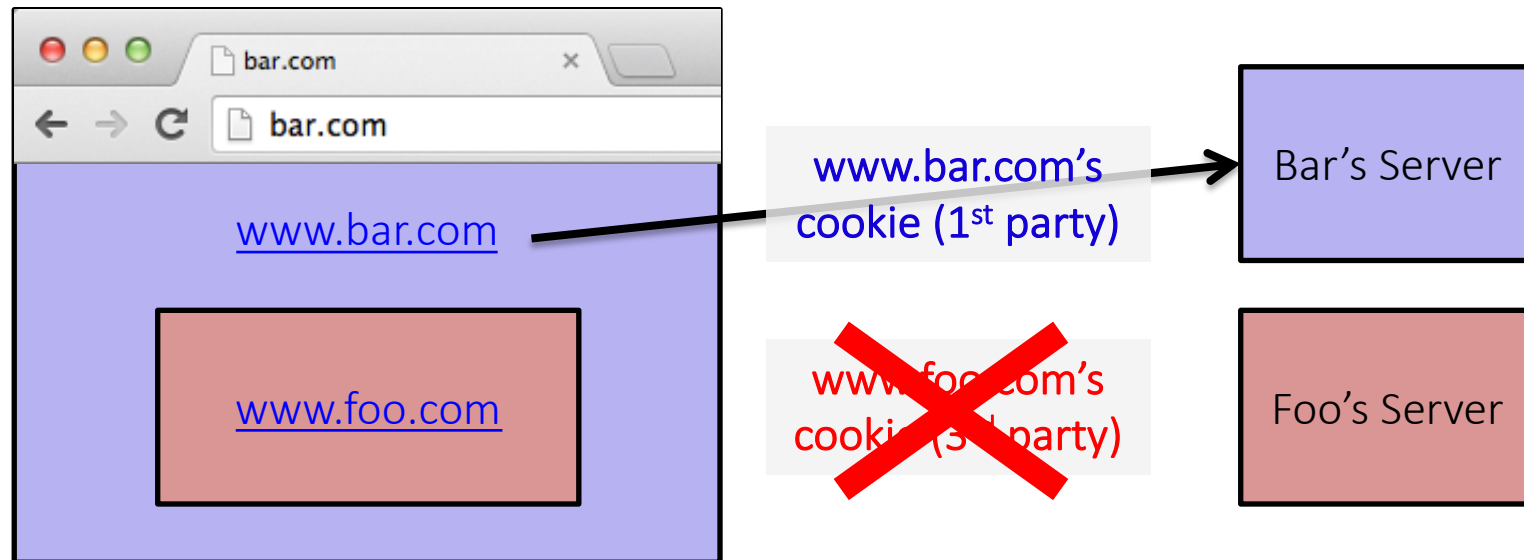
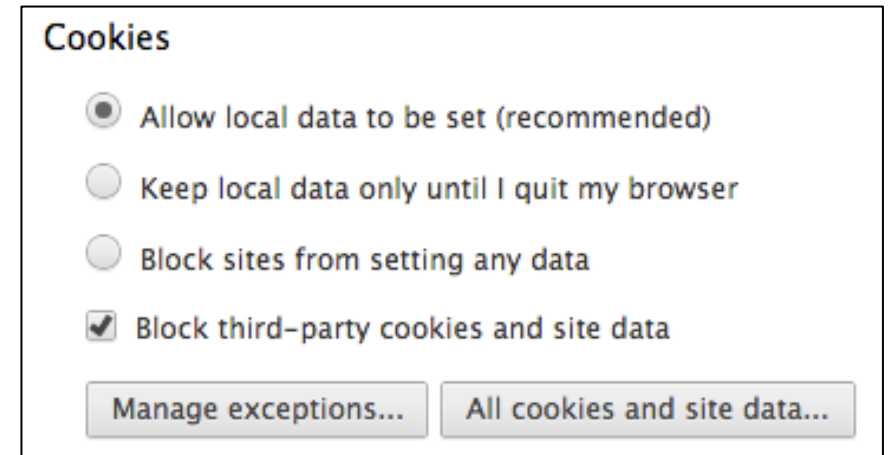
- Your browsing history
- Cookies and site data
- Information entered in forms

Your activity might still be visible to:

- Websites you visit
- Your employer or school
- Your internet service provider

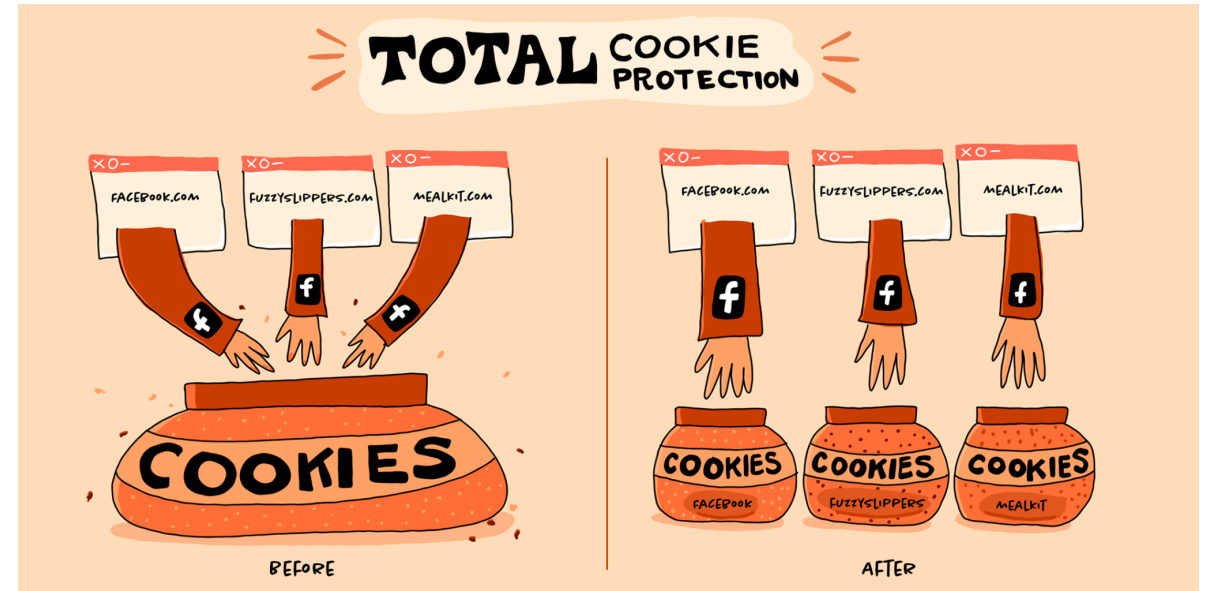
Defenses to Reduce Tracking

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



The Death of the Third-Party Cookie

- Essentially **no more 3rd party cookies** in Firefox or Safari
- Chrome is also moving towards **phasing out 3rd party cookies (2024)**
 - New proposal: “Topics”



There is still fingerprinting...

What Can Users Do?

- Choice of web browser



- Browser add-ons



uBlock Origin

