

CSE 484 / CSE M 584: Computer Security and Privacy

Web Security [Web Privacy]

Autumn 2020

Franziska (Franzi) Roesner
franzi@cs.washington.edu

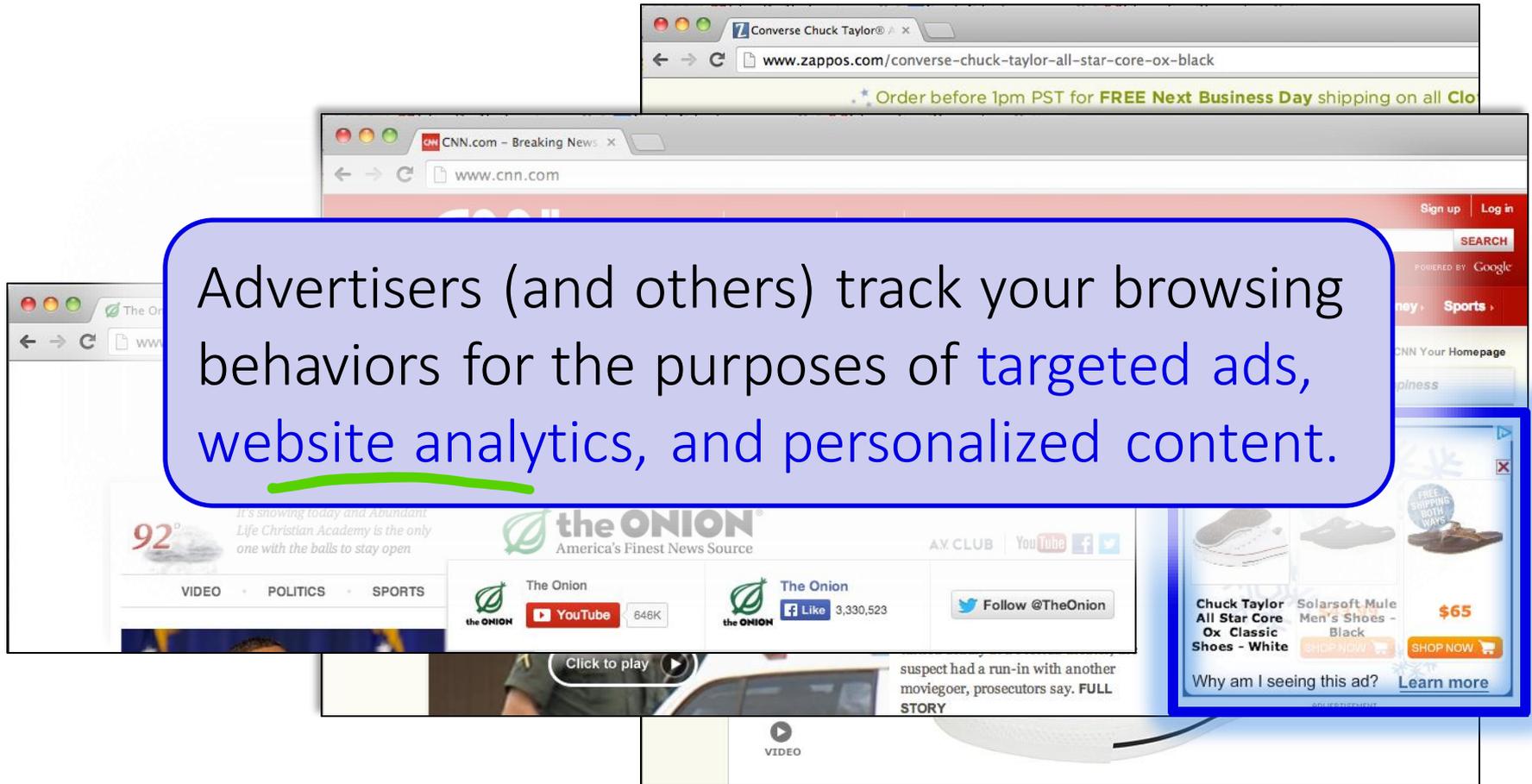
Thanks to Dan Boneh, Dieter Gollmann, Dan Halperin, Yoshi Kohno, Ada Lerner, John Manferdelli, John Mitchell, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials ...

Admin

- Lab 2:
 - Check access ASAP
 - Read FAQs 😊 .
- Homework 3: out soon, due 12/4
- Guest lecture on Friday
 - Emily McReynolds, law & policy
- No class the day before Thanksgiving .

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



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

Concerns About Privacy

The collage features several overlapping elements:

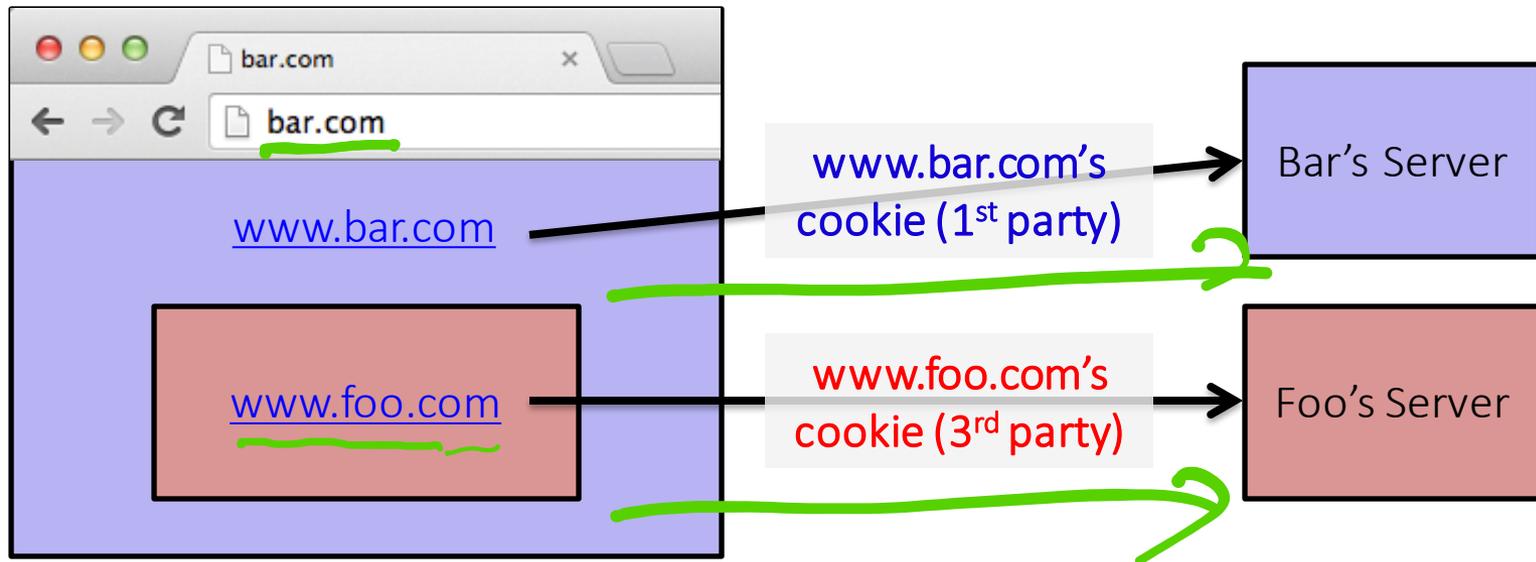
- THE WALL STREET JOURNAL.** (Top left)
- WHAT THEY KNOW | JULY 30, 2010** (Below WSJ)
- The New York Times** masthead (Center)
- May 6, 2011, 5:01 pm | 3 Comments** (Below NYT masthead)
- ‘Do Not Track’ Privacy Bill Appears in Congress** (Main headline)
- By TANZINA VEGA** (Byline)
- And the privacy legislation just keeps on coming.** (Sub-headline)
- 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.** (Lead paragraph)
- By JENNIFER VALENTINO-DEVRIES, JEREMY SINGER-VINE and ASHKAN SOLTANI** (Bottom byline)
- December 24, 2012** (Bottom date)
- CNN** logo (Left side)
- Log In** button (Right side)
- als ion** (Bottom right text)
- Big dep** (Left side text)
- Your Privacy** (Left side text)
- The file consists identifies her as** (Bottom left text)

Outline

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

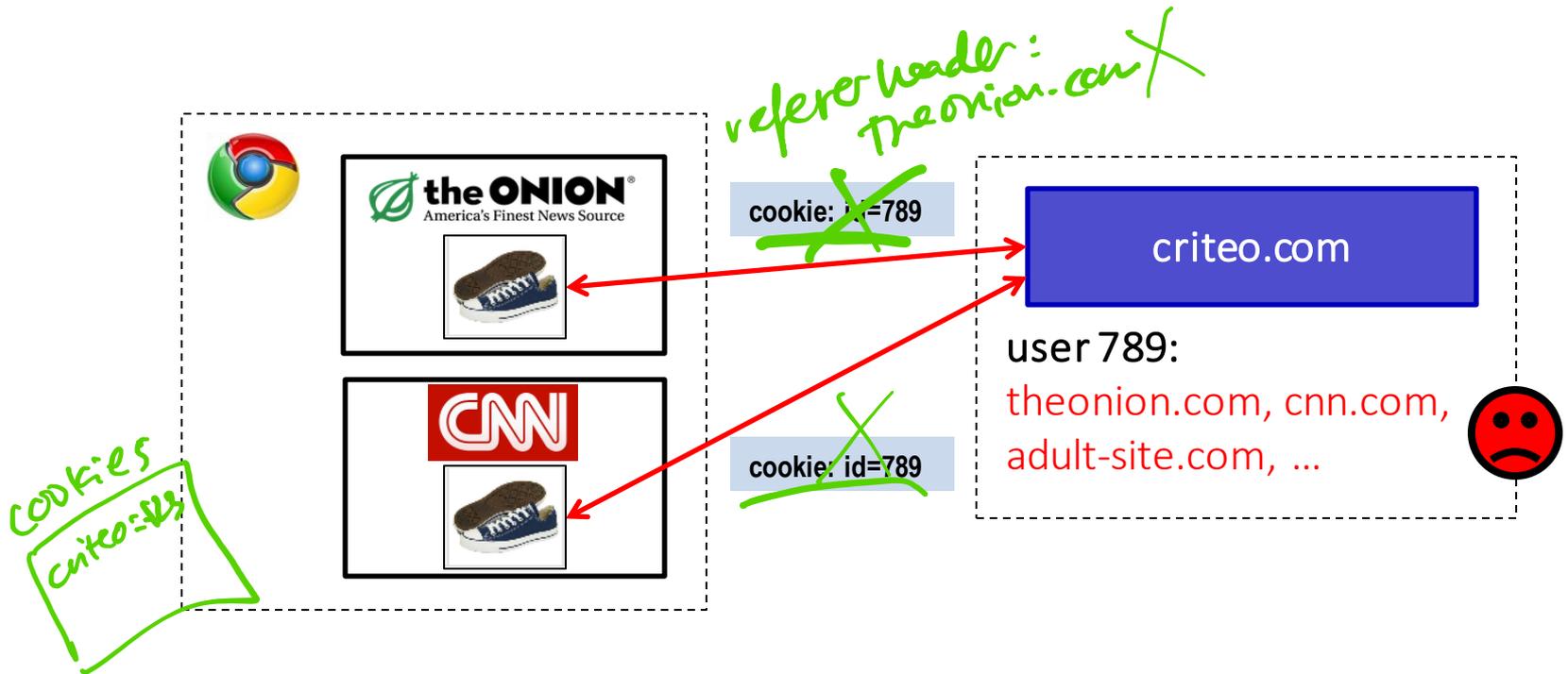
First and Third Parties

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



Anonymous Tracking

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



Basic Tracking Mechanisms

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

HTTP

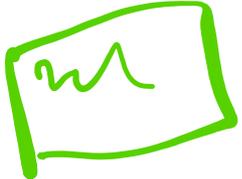
▼ Hypertext Transfer Protocol

```
▶ 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
```

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>)

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!) 



A research project of the [Electronic Frontier Foundation](#)

Panoptick

How Unique – and Trackable – Is Your Browser?

Is your browser configuration rare or unique? If so, web sites

Your browser fingerprint **appears to be unique** among the 3,435,834 tested so far

Only **anonymous data** will be collected by this site.



A paper reporting the statistical results of this experiment is now available: [How Unique Is Your Browser?](#), Proceedings of the Privacy Enhancing Technologies Symposium (PETS 2010), Springer Lecture Notes in Computer Science.

[Learn about Panoptick and web tracking.](#)

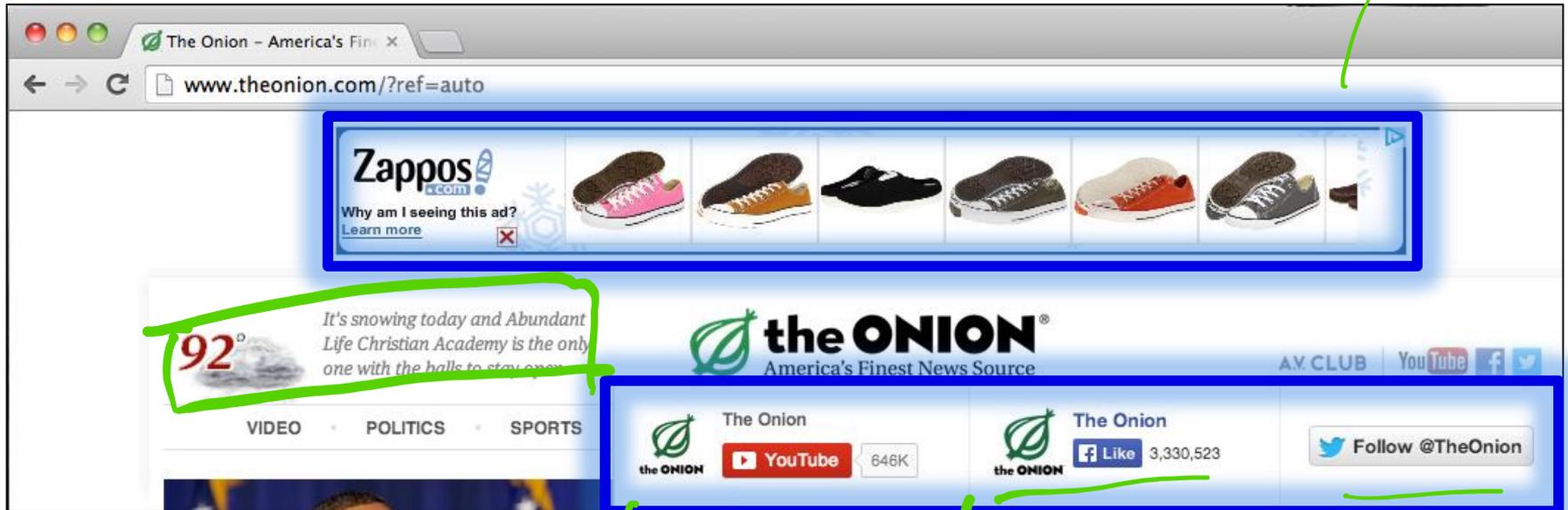
[The Panoptick Privacy Policy.](#)

[Learn about the Electronic Frontier Foundation.](#)

Brave

Other Trackers?

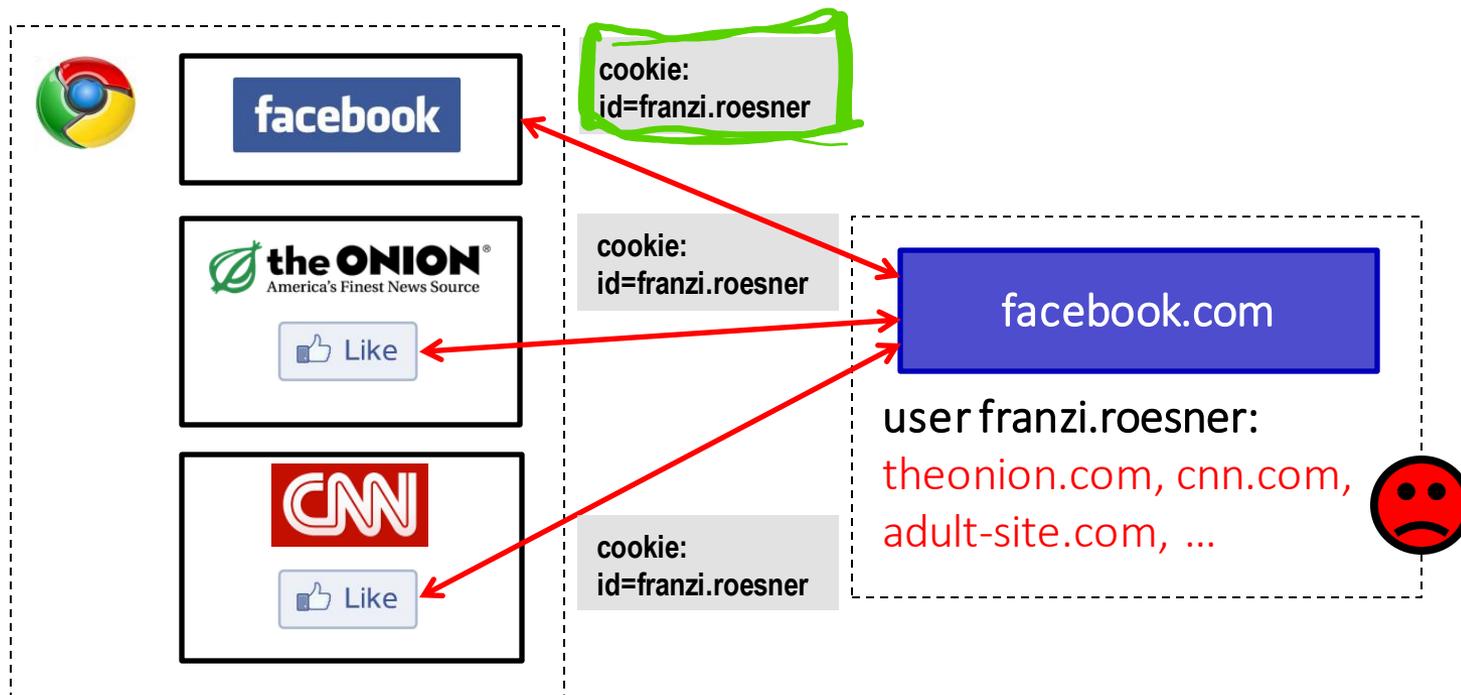
critico



“Personal” Trackers



Personal Tracking



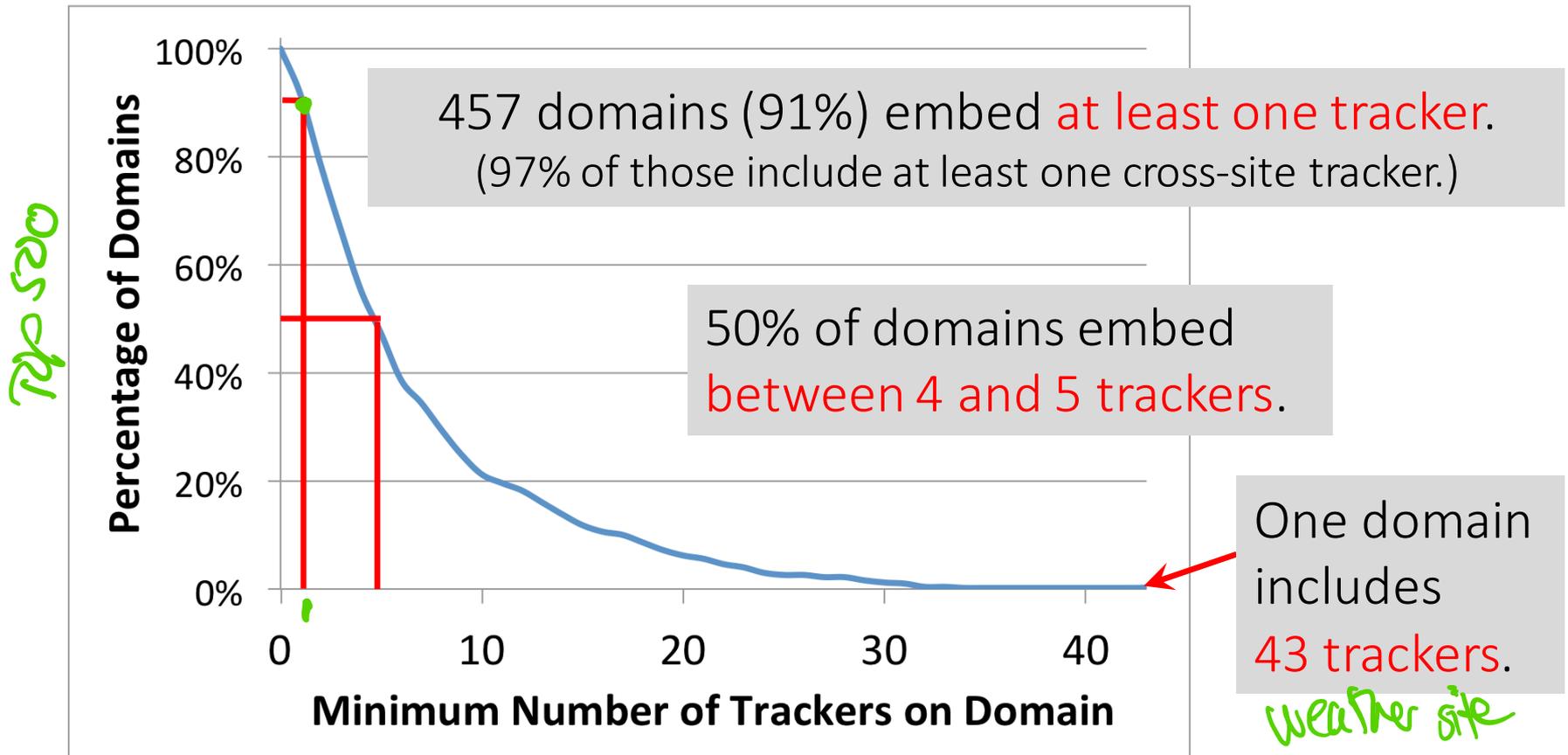
- Tracking is **not anonymous** (linked to accounts).
- Users **directly visit tracker's site** → evades some defenses.

Outline

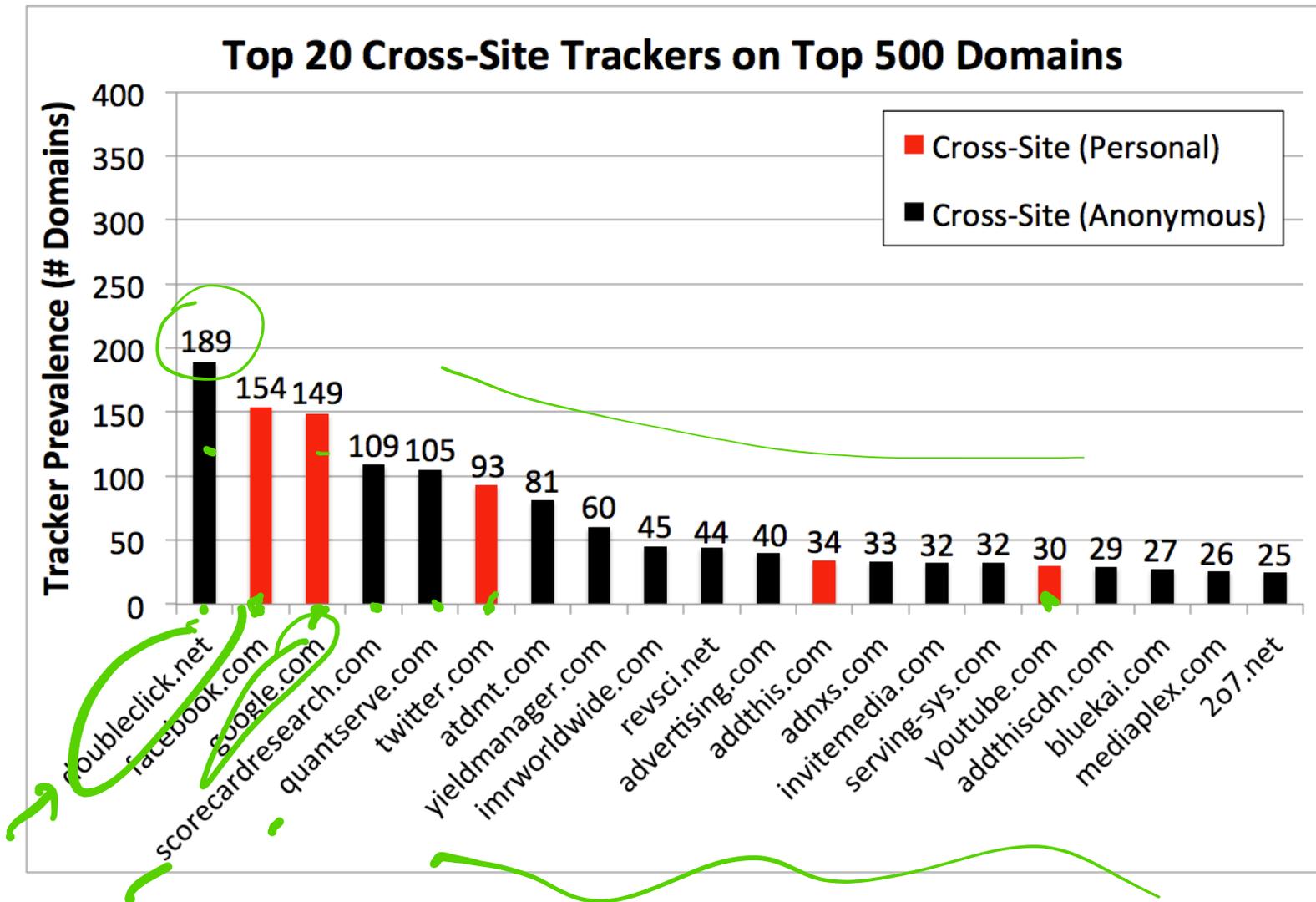
1. Understanding web tracking
2. Measuring web tracking
3. Defenses

How prevalent is tracking? (2011)

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



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? (our first study)
 - What about tracking before 2009? (first academic study)
- Solution: **time travel!**



The Wayback Machine to the Rescue

INTERNET ARCHIVE
Wayback Machine

1,260 captures
21 Dec 96 - 7 Oct 16

http://www.cs.washington.edu/ Go

NOV 1995 DEC 21 1996 JUN 1998

University of Washington
Computer Science & Engineering

Steam Powered Parallel Computing mural, circa 1986.

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](#).

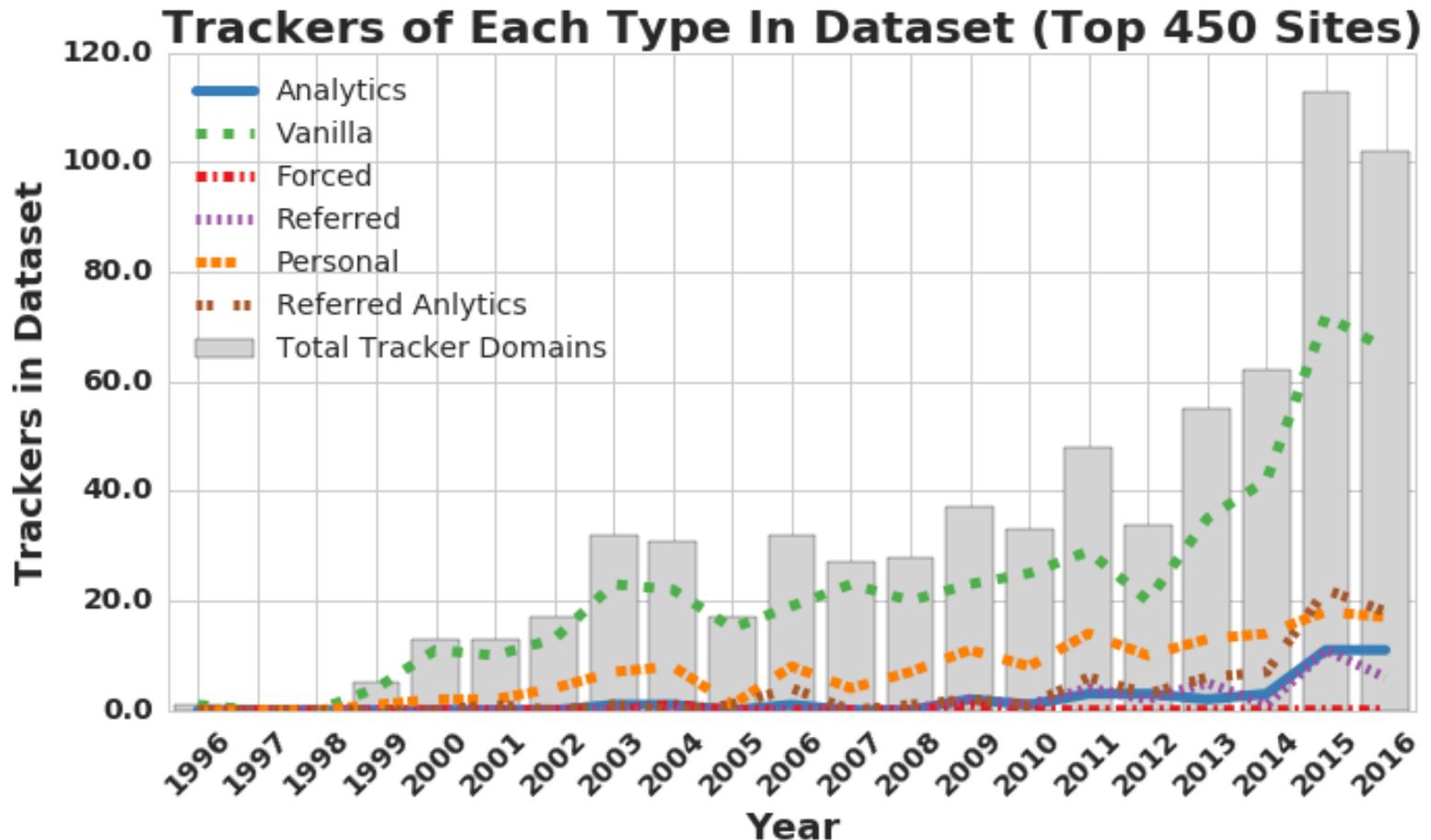
SPOTLIGHT

- **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>

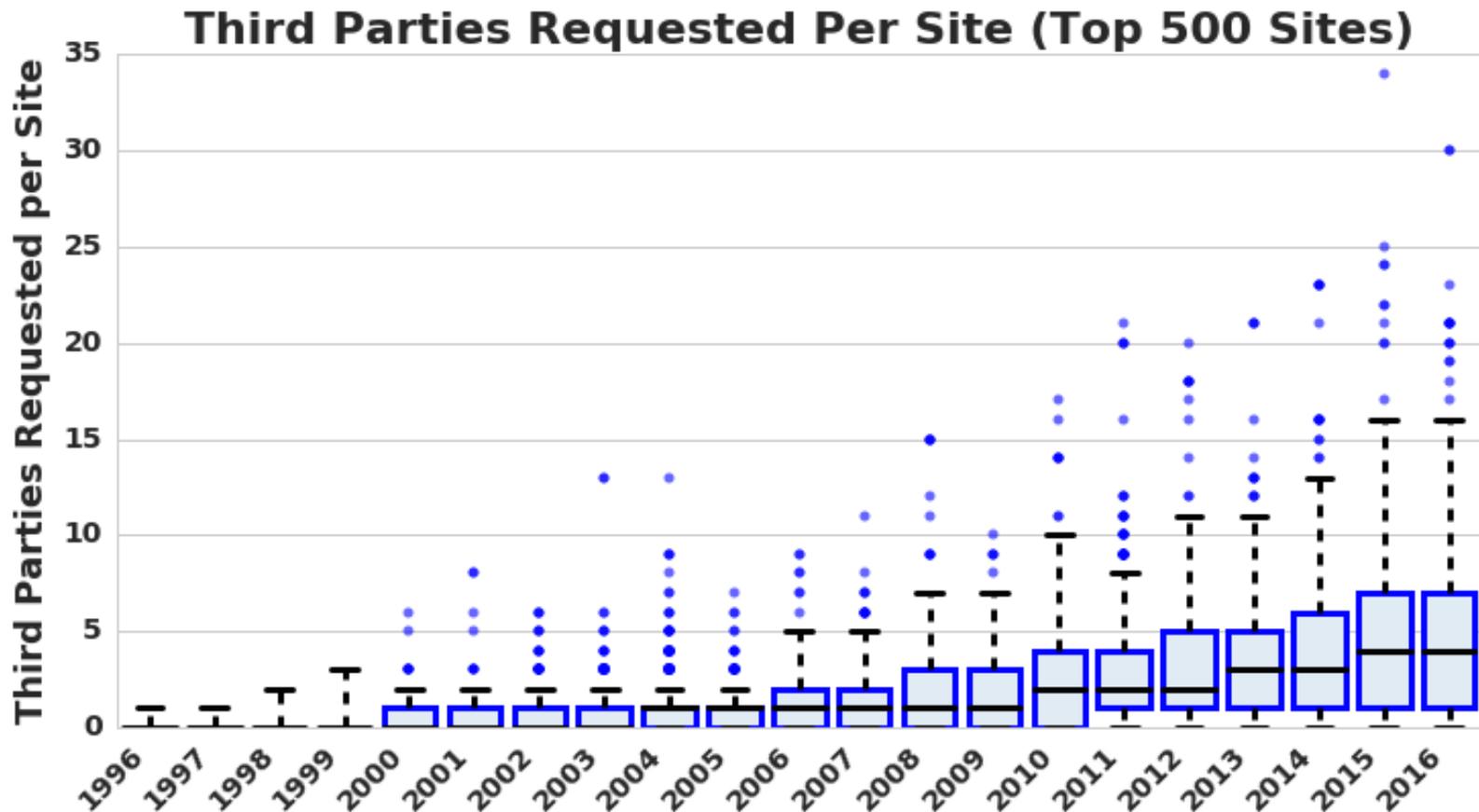
1996-2016: More & More Tracking

- More trackers of more types



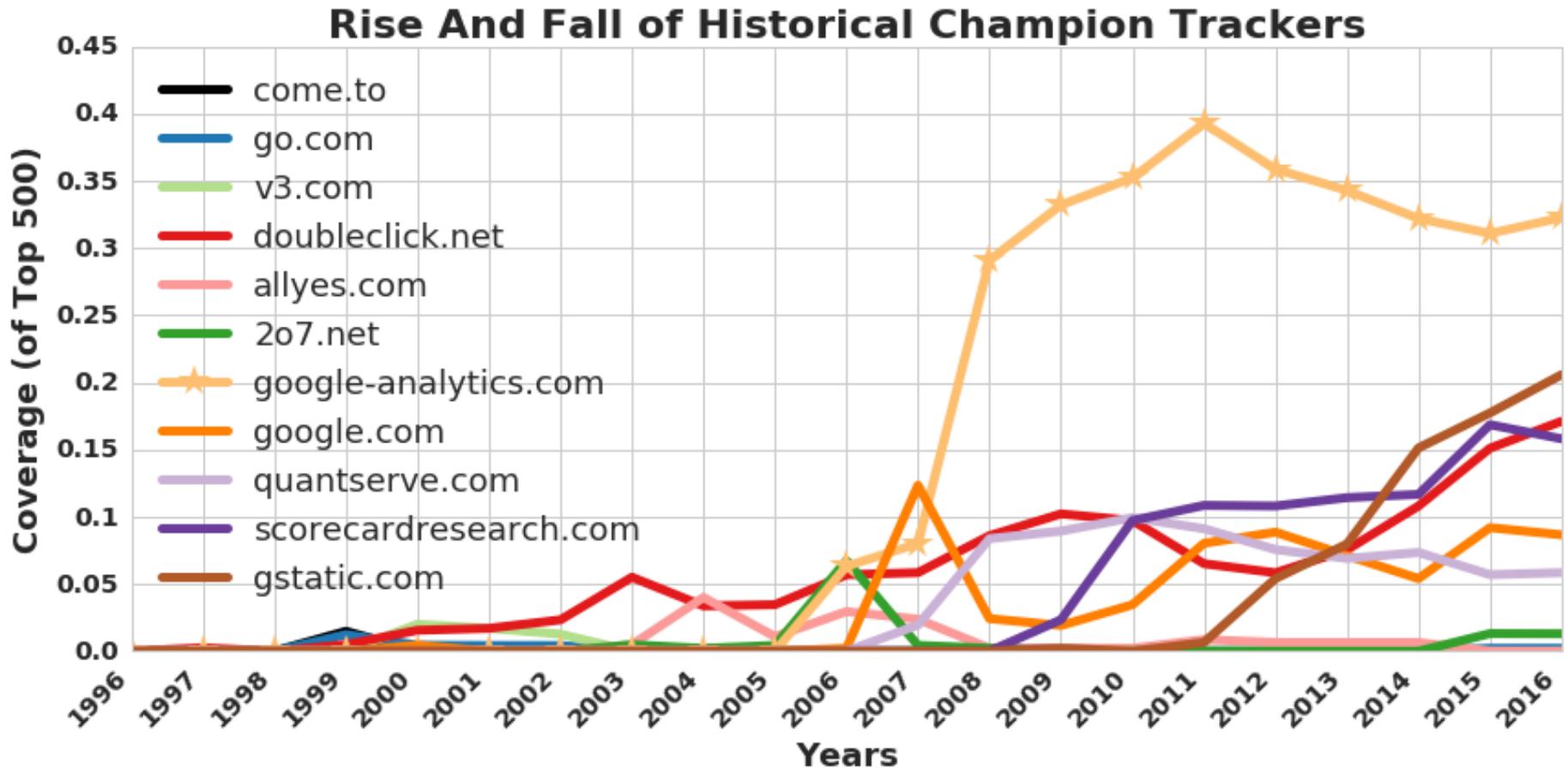
1996-2016: More & More Tracking

- 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?

DNT=1

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?

Private browsing mode 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:

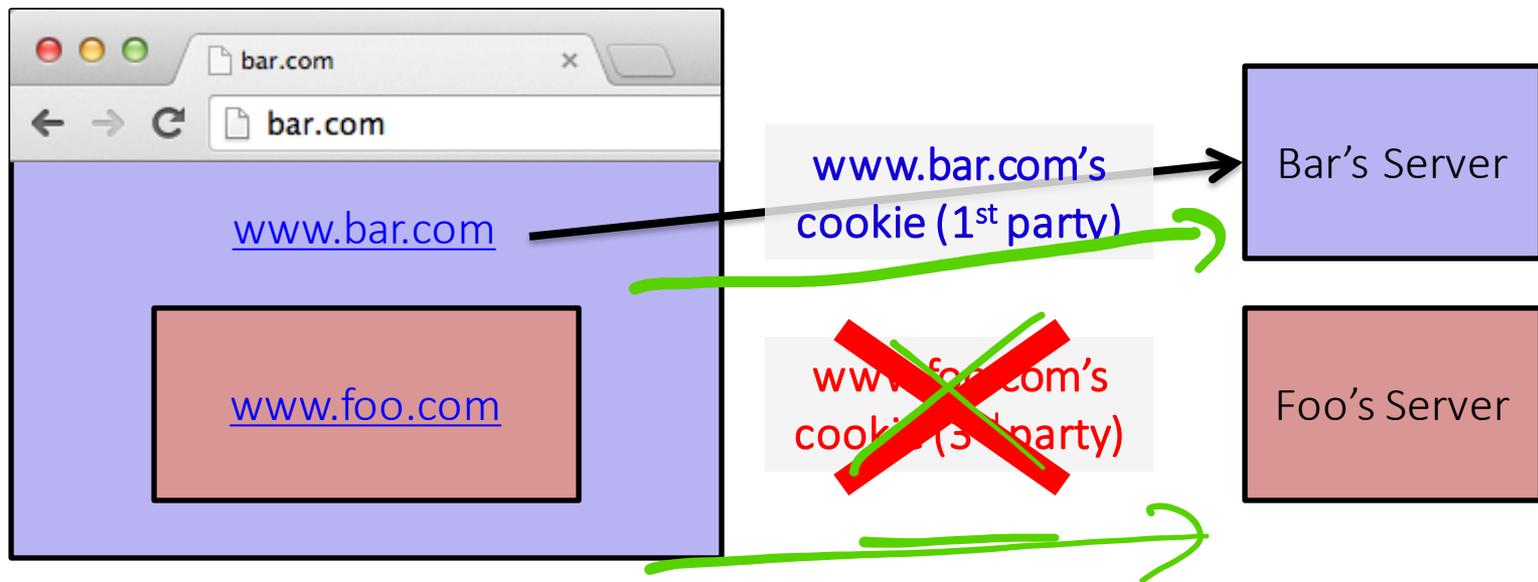
- Your browsing history
- Cookies and site data
- Information entered in forms

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?



Quirks of 3rd Party Cookie Blocking



In some browsers, this option means third-party cookies cannot be set, but **they CAN be sent.**

So if a third-party cookie is somehow set, **it can be used.**

How to get a cookie set?

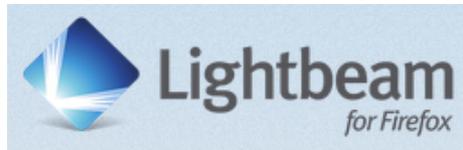
One way: be a first party.



etc.

Defenses to Reduce Tracking

- Do Not Track header?
- Private browsing mode?
- Third-party cookie blocking?
- Browser add-ons?



Often rely on blacklists, which may be incomplete.



“uses algorithmic methods to decide what is and isn't tracking”

