

# Browser fingerprinting

Nataliia Bielova

[@nataliabelova](https://twitter.com/nataliabelova)

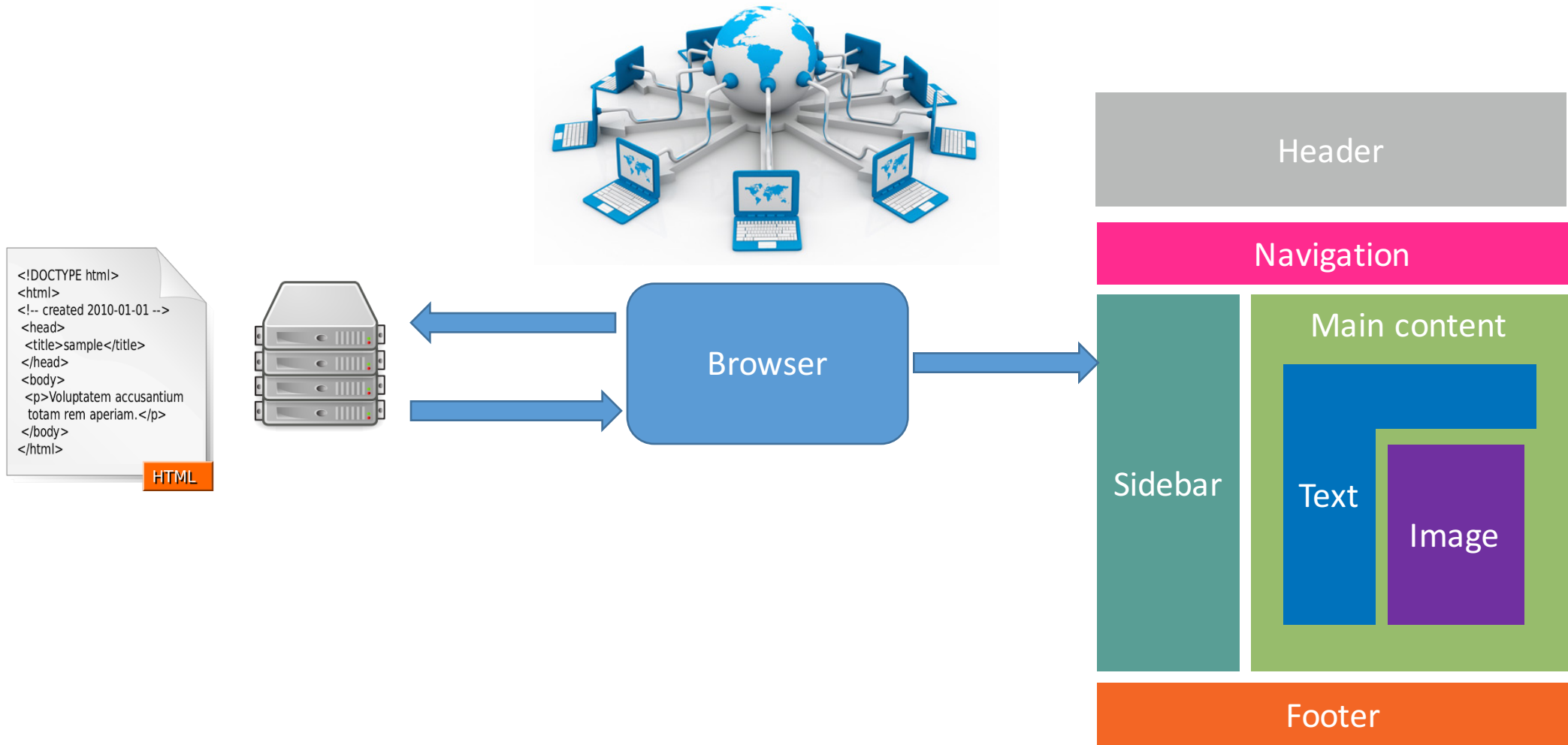
February 12<sup>th</sup>, 2019

Security and ethical aspects of data

Université Côte d'Azur

# Today's class

- A brief history of Web browsers
- What is browser fingerprinting?
- From basic to advanced fingerprinting



# I. Internet in 1995

## HTTP User agent

NCSA\_Mosaic/2.0  
(Windows 3.1)

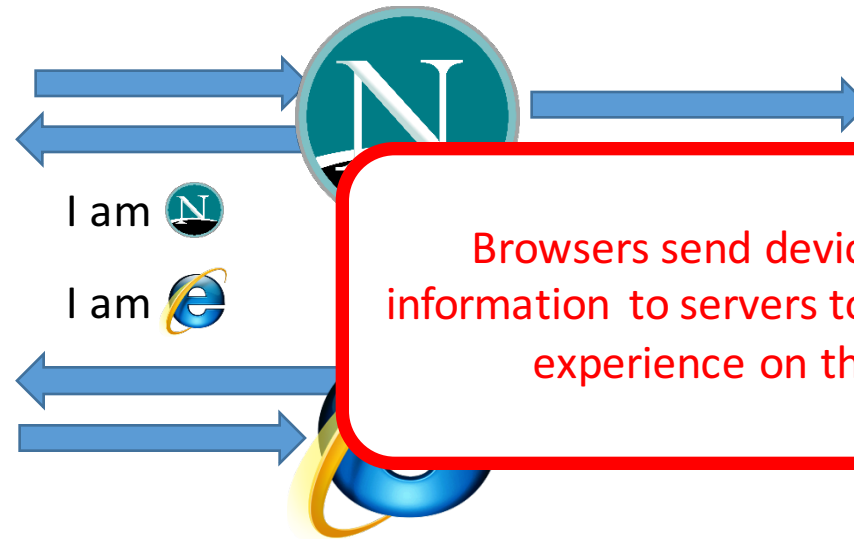
```
<!DOCTYPE html>
<html>
<!-- created 2010-01-01 -->
<head>
<title>sample</title>
</head>
<body>
<p>Voluptatem accusantium
totam rem aperiam.</p>
</body>
</html>
```

HTML

Mozilla/1.22  
(compatible; MSIE  
2.0; Windows 95)

```
<!DOCTYPE html>
<html>
<!-- created 2010-01-01 -->
<head>
<title>sample</title>
</head>
<body>
<p>Voluptatem accusantium
totam rem aperiam.</p>
</body>
</html>
```

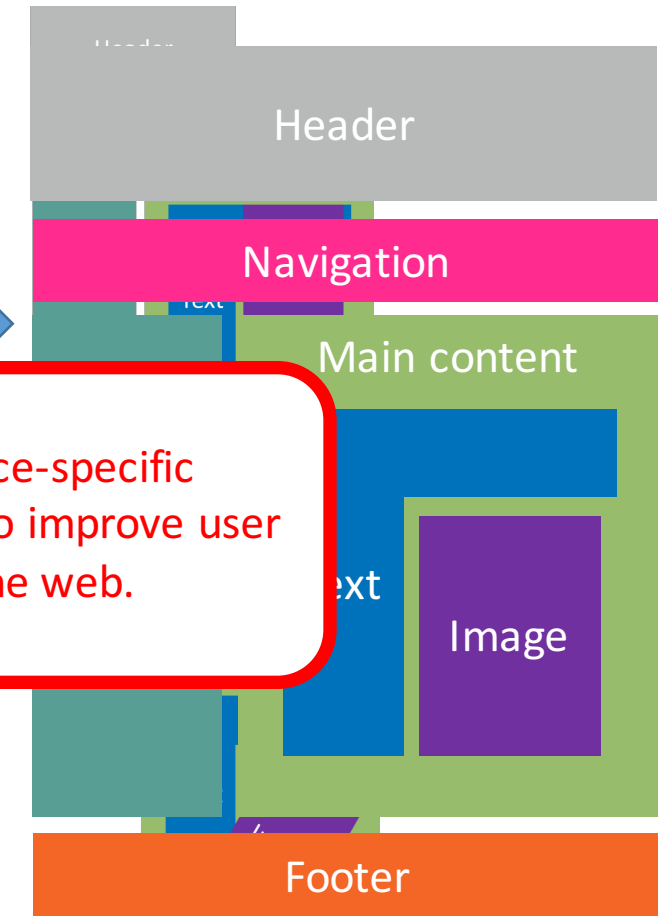
HTML



Browsers send device-specific information to servers to improve user experience on the web.

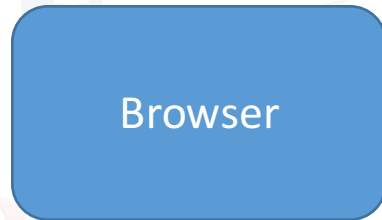
```
<!DOCTYPE html>
<html>
<!-- created 2010-01-01 -->
<head>
<title>sample</title>
</head>
<body>
<p>Voluptatem accusantium
totam rem aperiam.</p>
</body>
</html>
```

HTML



- Every website announces with **what browser** it is recommended to visit the website





1995	2017
Browser: Netscape Language: Fr	Browser: Chrome v53 OS: Linux Screen: 1920x1080 Language: Fr Timezone: GMT+1 Graphic card: GTX 1080Ti ...

A bigger and richer web



- Audio
- Video
- 3D rendering
- Real-time communications
- Web payments
- Virtual reality

What happens when we start collecting all the information available in a web browser?

# Example of a browser fingerprint

Attribute	Value
User agent	Mozilla/5.0 (X11; Fedora; Linux x86_64; rv:55.0) Gecko/20100101 Firefox/55.0
HTTP headers	text/html, application/xhtml+xml, application/xml;q=0.9,*/*;q=0.8 gzip, deflate, br en-US,en;q=0.5
Plugins	Plugin 0: QuickTime Plug-in 7.6.6; libtotem-narrow-space-plugin.so; Plugin 1: Shockwave Flash; Shockwave Flash 26.0 r0; libflashplayer.so.
Fonts	Century Schoolbook, Source Sans Pro Light, DejaVu Sans Mono, Bitstream Vera Serif, URW Palladio L, Bitstream Vera Sans Mono, Bitstream Vera Sans, ...
Platform	Linux x86_64
Screen resolution	1920x1080x24
Timezone	-480 (UTC+8)
OS	Linux 3.14.3-200.fc20.x86_64 32-bit
WebGL vendor	NVIDIA Corporation
WebGL renderer	GeForce GTX 650 Ti/PCIe/SSE2
Canvas	<p>Cwm fjordbank glyphs vext quiz, ☺</p> <p>Cwm fjordbank glyphs vext quiz, ☺</p>



Maverick  
Ocean Front Villas  
Mandarin Tea  
Regency  
Sassafras & Ginger  
Dollhouse  
Athletics Dept.



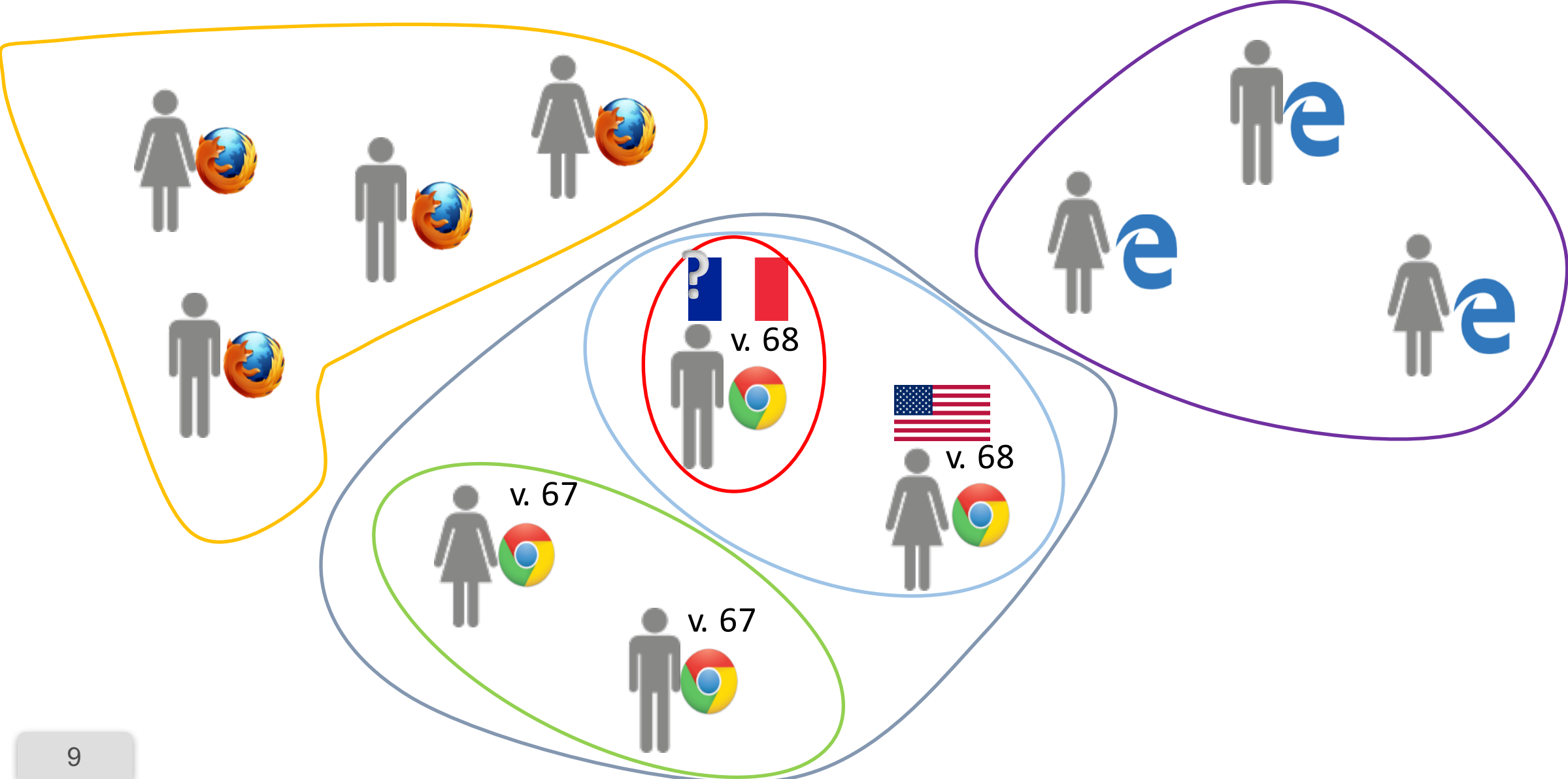
# I. Definition of browser fingerprinting

## Definitions

- A **browser fingerprint** is a set of information related to a user's device from the hardware to the operating system to the browser and its configuration.
- Browser **fingerprinting** refers to the process of collecting information through a web browser to build a fingerprint of a device.

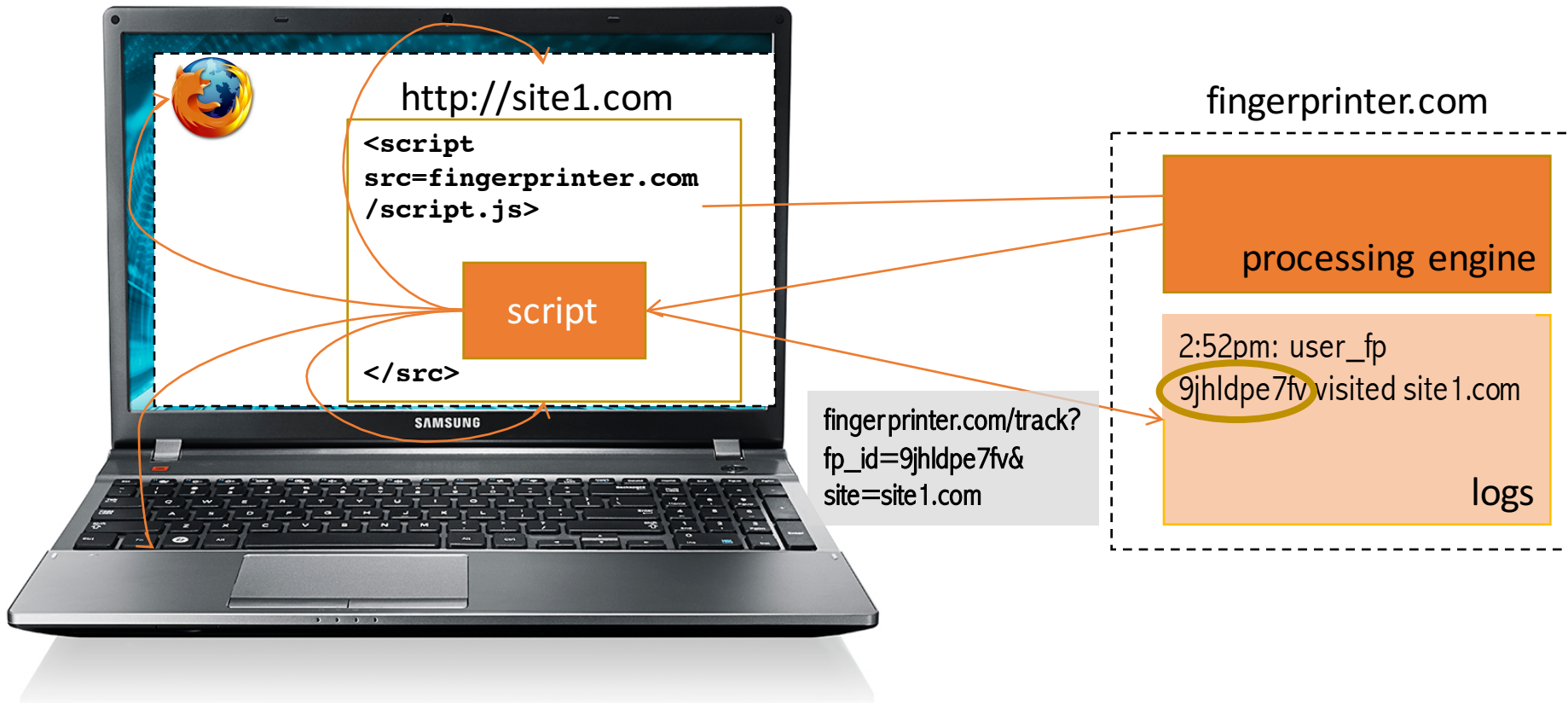


# How can we be identified by a browser fingerprint?



# Browser fingerprinting used for tracking

**Browser and operating system properties** are used to **track repeated visits** to a site.



# Comparison of the emoji on different devices and OSs



(a) Windows 7



(b) Windows 10



(c) Linux



(d) iOS



(e) Firefox OS



(f) Android 4.3 and before



(g) Android 4.4



(h) Android 5.0



(i) Android on an LG device



(j) Android on a Samsung device



(k) Android on an HTC device



(l) Emoji not supported

<https://hal.inria.fr/hal-01285470/document>

Two studies have investigated the diversity of browser fingerprints.



Am I Unique?

470,161 fingerprints  
94.2% were unique

118,934 fingerprints  
89.4% were unique

Tracking is possible

# Fingerprinting



- Panopticlick [Eckersley, PET'2010]

Your browser fingerprint **appears to be unique** among the 2,419,678 tested so far.

Currently, we estimate that your browser has a fingerprint that conveys **at least 21.21 bits of identifying information.**

- Information needed to **uniquely identify a browser**
  - $n$  – number of connected devices: **5 000 000 000**
  - $\log_2 n$  – number of bits for a unique id: **33 bits**
- **Idea: distinguish user's browsers** by accessing browser features and using their probability distributions



## <https://amiunique.org> (Am I Unique)

The screenshot shows the homepage of the Am I Unique website. It features a teal header with the site name, a dark sidebar with navigation links, and a main content area with a central call-to-action button and a privacy notice.

Am I Unique?

- Home
- My fingerprint
- Global statistics
- FAQ
- Privacy policy
- Links
- About
- View on GitHub

Learn how identifiable you are on the Internet  
Help us investigate the diversity of web browsers

[View my browser fingerprint](#)

By clicking on this button, only anonymous data will be collected and a cookie will be stored in your browser for four months. You can find more details in the [Privacy Policy](#).

Spread the word! Share AmlUnique!  
Try it on all your devices!

[What is browser fingerprinting?](#) [Learn more](#)

Any questions? Send us an email at [contact@amiunique.org](mailto:contact@amiunique.org)

- Website launched in November 2014
- Collected 660,000+ fingerprints so far
- Browser extension available to see the evolution of your own browser fingerprint

Cwm fjordbank glyphs vext quiz, 😊

Cwm fjordbank glyphs vext quiz, 😊

# How to compare datasets: Anonymity sets

- User-agent on Desktop vs Mobile devices

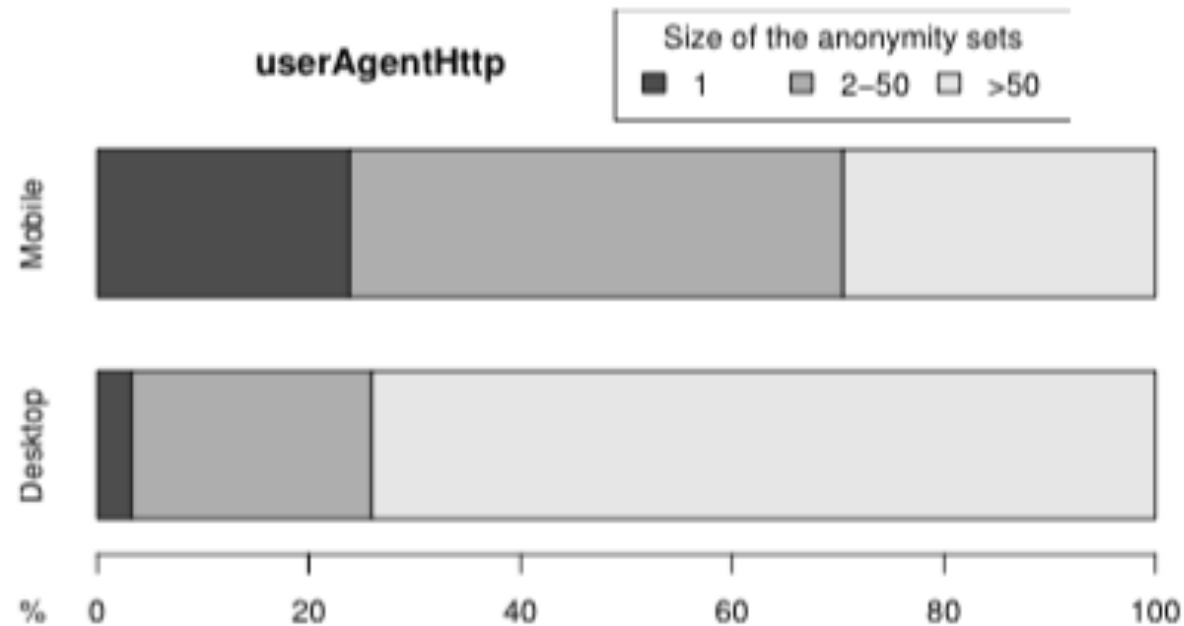


Fig. 4. Comparison of anonymity set sizes on the user-agent between desktop and mobile devices



# I. Example of values collected on AmlUnique

## Some user-agents

- Mozilla/5.0 (Windows NT 6.1; WOW64; rv:34.0) Gecko/20100101 Firefox/34.0
- Mozilla/5.0 (iPhone; CPU iPhone OS 8\_1\_2 like Mac OS X) AppleWebKit/600.1.4 (KHTML, like Gecko) Version/8.0 Mobile/12B440 Safari/600.1.4
- Mozilla/5.0 (Android; Mobile; rv:27.0) Gecko/27.0 Firefox/27.0
- Mozilla/5.0 (Macintosh; Intel Mac OS X 10\_10\_2) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/39.0.2171.95 Safari/537.36
- Mozilla/5.0 (X11; Ubuntu; Linux i686; rv:34.0) Gecko/20100101 Firefox/34.0

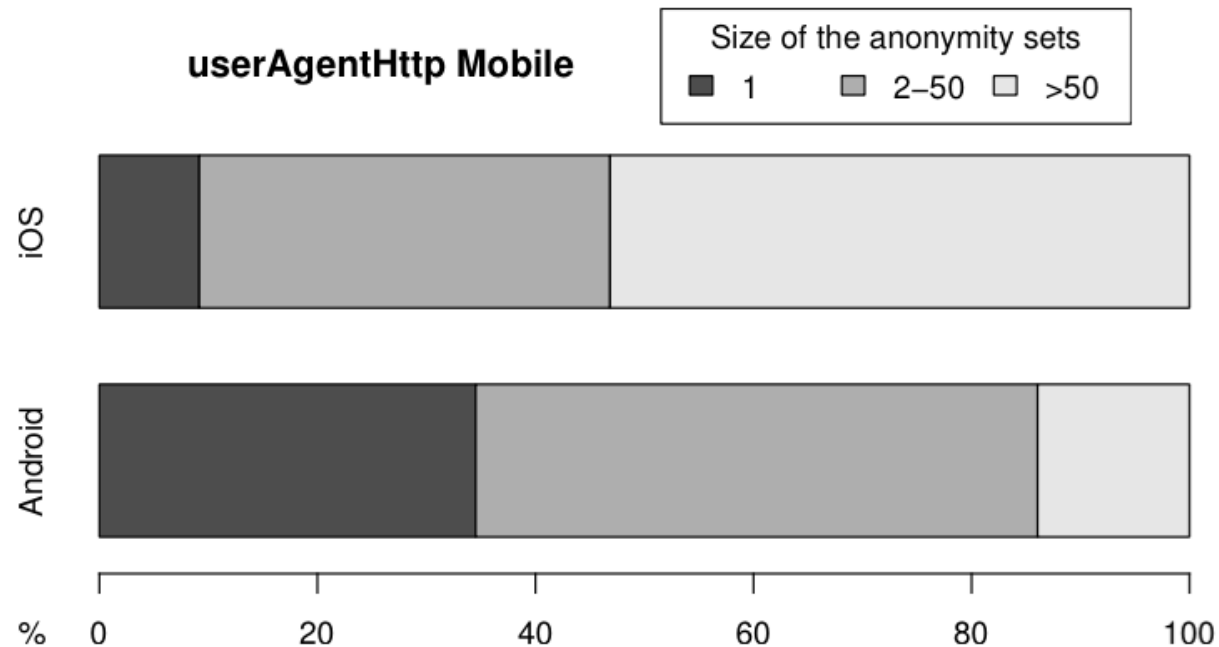
# I. Example of values collected on AmlUnique

## Other custom user-agents

- godzilla/5.0 (X122; BSD; rv:500.0) Gecko/20100101
- pouet
- “54. When a warlike prince attacks a powerful state, his generalship shows itself in preventing the concentration of the enemy's forces. He overawes his opponents, and their allies are prevented from joining against him.”
- Deepnet Explorer 1.5.3; Smart 2x2; Avant Browser; .NET CLR 2.0.50727; InfoPath.1)
- NSA
- Game Boy Advance
- eat it

# Anonymity sets for mobile devices

- User-agent on Android vs iOS devices



› Fig. 5. Comparison of anonymity set sizes on the user-agent between Android and iOS devices

# What if I disable JavaScript?

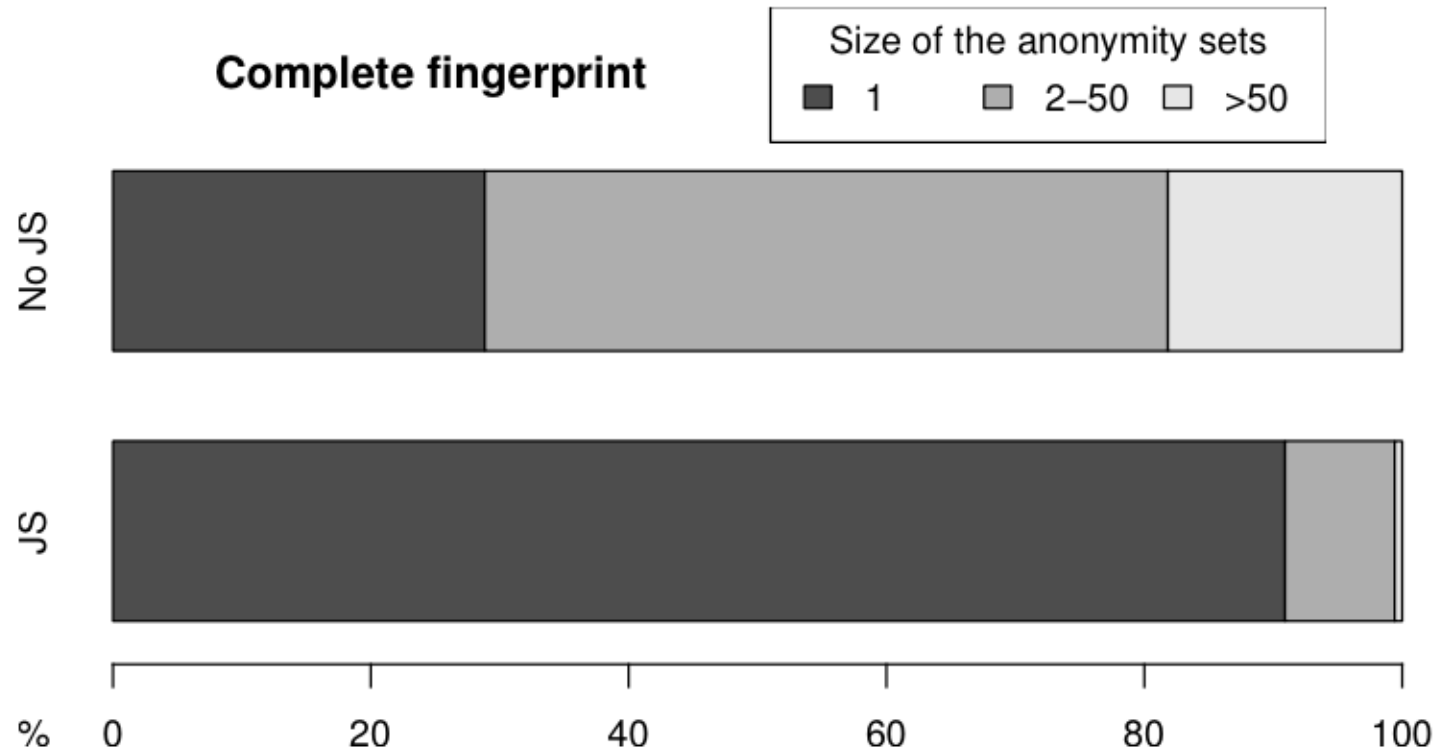


Fig. 9. Comparison of anonymity set sizes on the complete fingerprint between devices with and without JavaScript

- Servers can easily collect information about a device to form what is called a **browser fingerprint**.
- There is so much diversity that users can be **tracked** online if their fingerprint is **unique**.
- Test your device on <https://amiunique.org> and <https://extensions.inrialpes.fr>

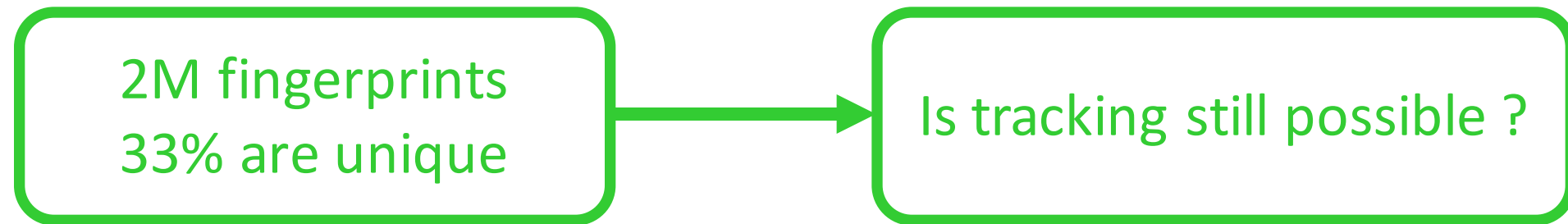
# Very hard to opt-out

- Even if
  - you delete all the cookies
  - you clean all the storages (HTML5, Flash)
  - you use browser private mode

...your fingerprint remains the same!



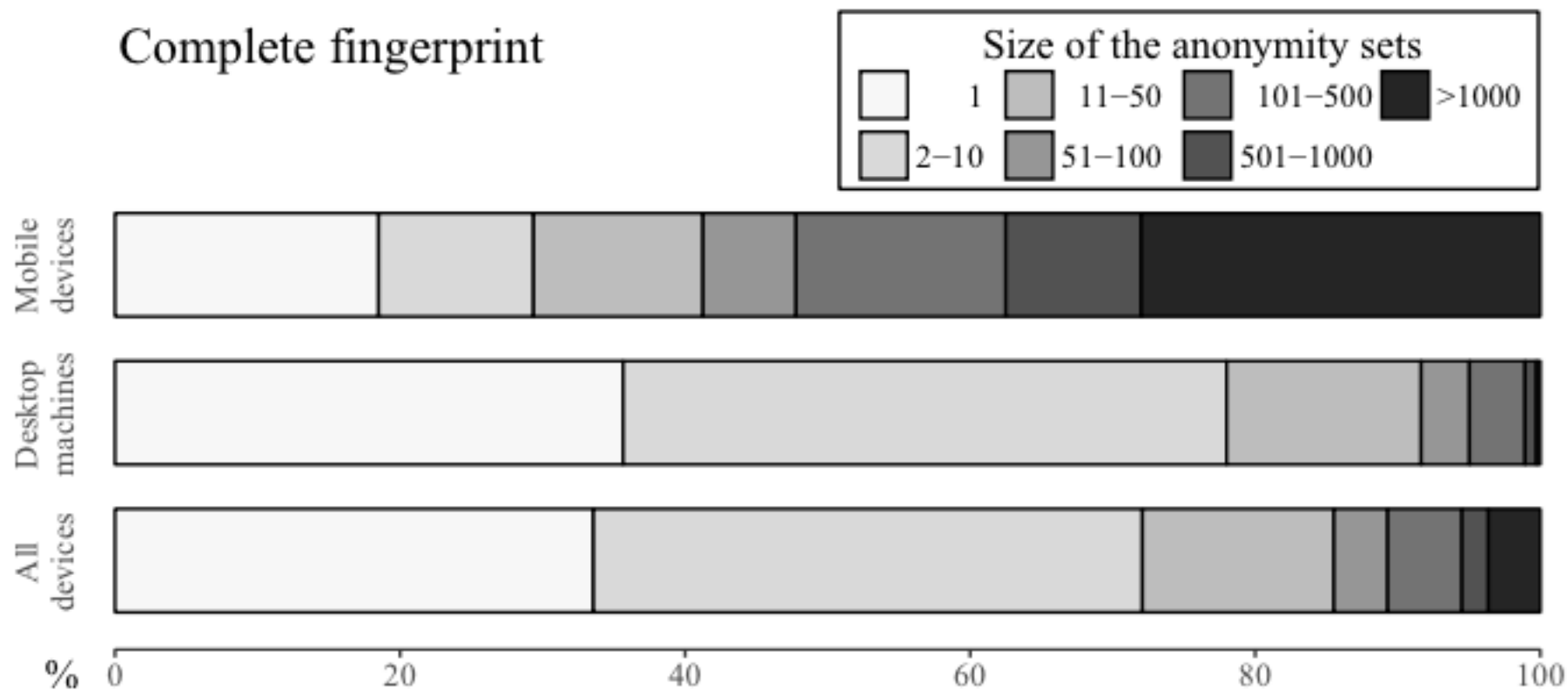
- How effective is fingerprinting at large scale?



## **Hiding in the Crowd: an Analysis of the Effectiveness of Browser Fingerprinting at Large Scale**

*Alejandro Gómez-Boix, Pierre Laperdrix, Benoit Baudry*

The Web Conference ([WWW 2018](#))



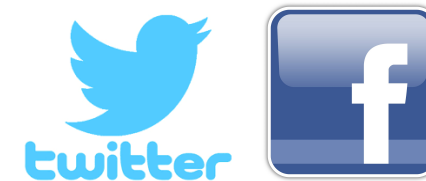
**Figure 3: Comparison of anonymity set sizes between mobile devices and desktop/laptop machines.**



# New Fingerprinting Methods

- **Privacy Paradox**

- Users' fingerprints can be enriched by their browser extensions
- Moreover, we found an attack allows to detect 58 web services where the user is logged in!



# I. Plugins VS Browser extensions

- **Plugins** were created to display content not supported by the browser

- Flash    Java    Silverlight



- All installed plugins are accessible via the `navigator.plugins` JavaScript object

- **Extensions** extend or modify default behavior of a browser

- AdBlockPlus, LastPass, Ghostery, Pinterest



LastPass...



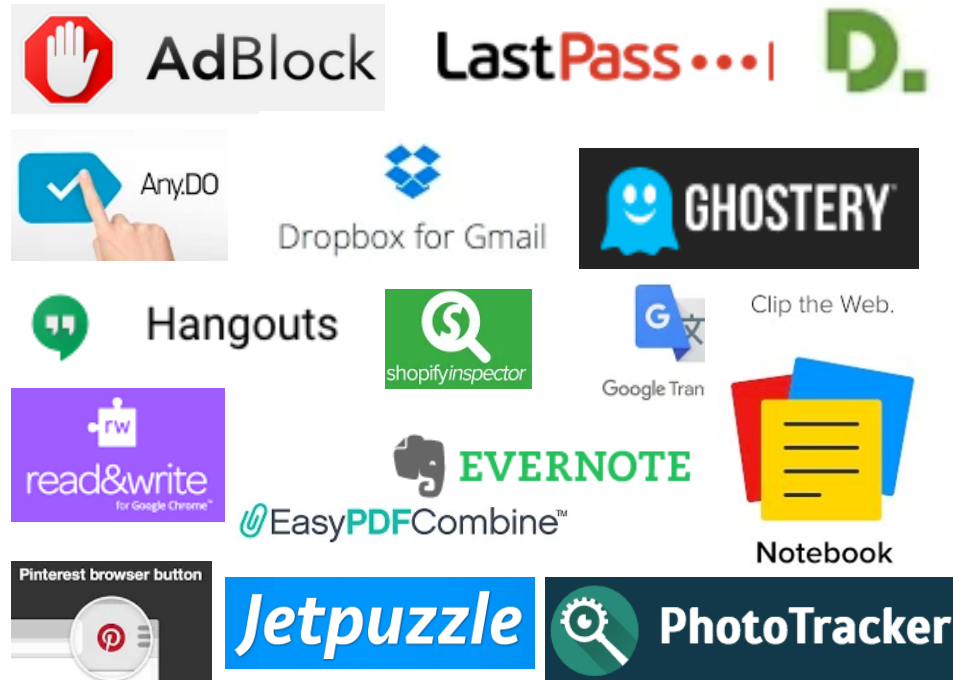
- There is no API that webpages can use to detect all installed extensions



# How unique is your browser?

<https://extensions.inrialpes.fr>

- Browser extension detection
- ~13 000 extensions



- Websites a user is logged in
- 58 websites



# Browser extension detection

- via **Web Accessible Resources**

[chrome-extension://gpdjojdkbbmdfjfahjcgigfpmkopogic/img/icon\\_48.png](chrome-extension://gpdjojdkbbmdfjfahjcgigfpmkopogic/img/icon_48.png)

unique extension ID

## Discovering Browser Extensions via Web Accessible Resources

Alexander Sjösten  
Chalmers University of  
Technology  
Gothenburg, Sweden  
sjosten@chalmers.se

Steven Van Acker  
Chalmers University of  
Technology  
Gothenburg, Sweden  
acker@chalmers.se

Andrei Sabelfeld  
Chalmers University of  
Technology  
Gothenburg, Sweden  
andrei@chalmers.se

### ABSTRACT

Browser extensions provide a powerful platform to enrich browsing experience. At the same time, they raise important security questions. From the point of view of a website, some browser extensions are invasive, removing intended features and adding unintended ones, e.g. extensions that hijack Facebook likes. Conversely, from the point of view of extensions, some websites are invasive, e.g. websites that bypass ad blockers. Motivated by security goals at clash, this

The first and second scenarios present an exclusive point of view of websites, concerned with malicious extensions. The third scenario presents an exclusive view of extensions, concerned with malicious websites. The fourth scenario illustrates legitimate synergies between websites and extensions. Finally, the fifth scenario illustrates the security goals of websites and extensions at outright clash.

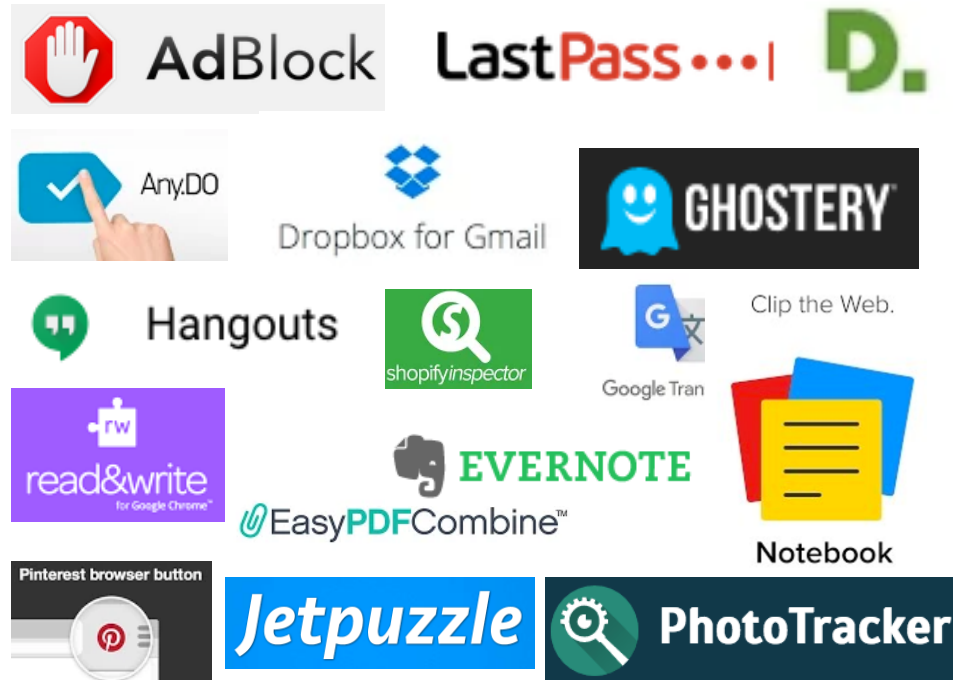
**Bank scenario** Bank webpages manipulate sensitive information whose unauthorized access may lead to financial



# How unique is your browser?

<https://extensions.inrialpes.fr>

- Browser extension detection
- ~13 000 extensions



- Websites a user is logged in
- 58 websites



# Detection of websites a user logged in

- Redirection URL hijacking [@robin\\_linus](#)
- Abusing Content Security Policy (CSP) – no JavaScript needed [@homakov](#)

## Your Social Media Fingerprint

Without your consent most major web platforms leak whether you are logged in. This allows any website to detect on which platforms you're signed up. Since there are lots of platforms with specific demographics an attacker could reason about your personality, too.

This project is an open source contribution of [RobinLinus - Security, Privacy & Blockchain Consulting](#).

### Demonstration

You are logged in to:



Monday, January 13, 2014

## Using Content-Security-Policy for Evil

**TL;DR** How can we use technique created to protect websites for Evil? (We used [XSS Auditor](#) for Evil before) There's a neat way: taking advantage of CSP we can detect whether URL1 does redirect to URL2 and even bruteforce /path of URL2/path. This is a conceptual vulnerability in CSP design (violation == detection), and there's no obvious way to fix it.

Demo & playground: <http://homakov.github.io/csp.html>

# How unique is your browser?

<https://extensions.inrialpes.fr>

## Browser Extension and Login-Leak Experiment

When you browse the web, **small beacons** (trackers) are spying on your online activities. Even though such trackers are invisible, they collect information about you such as which pages you visit, which buttons clicked, and what text you typed. This information is often used to show you **targeted advertisements** and **may require you to pay a higher price during online shopping** depending on the collected information.

Did you know websites can track you by your browser extensions and web logins?

Recent studies show that you can be tracked **based on your web browser properties**. In this experiment, we demonstrate that you can also be tracked by

- your browser extensions (such as AdBlock, Pinterest, or Ghostery), and
- the websites you have logged in (such as Facebook, Gmail, or Twitter).

You can learn more here about how these detection techniques work.

In the experiment, we will collect your browser fingerprint, together with the browser extensions installed and a list of websites you have logged in. We only collect anonymous data during the experiment (see our **Privacy Policy**), we will securely store the data on an Inria server, use it only for research purpose and not share it with anyone outside of Inria. You can also read **the frequently asked questions here**.

**21 000 users  
have already tested!**

Browser will silently visit **these sites**.

(we would like to see whether our dataset is biased)  
Regular computer user.  I don't want to declare.

I agree, test my browser!

# How unique is your browser?

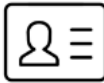

<https://extensions.inrialpes.fr>

← → ↻ <https://extensions.inrialpes.fr> ★

Main page News & Updates How... ∨ F.A.Q. Privacy policy English

## Are you identifiable?

**Yes, you are identifiable**, as there are no other users who looks like you among the 21939 users we tested so far:

 ← Easily trackable... More anonymity... →  More similar users

None 1 2 3 4 5 6 7 8 9+ You are here

### Are you identifiable...

- ...by your **extensions**? **no**
- ...by your **website logins**? **no**
- ...by your **browser fingerprint**? **no**
- ...by your **extensions, web logins and browser fingerprint together**? **yes**



# User dataset w.r.t previous studies

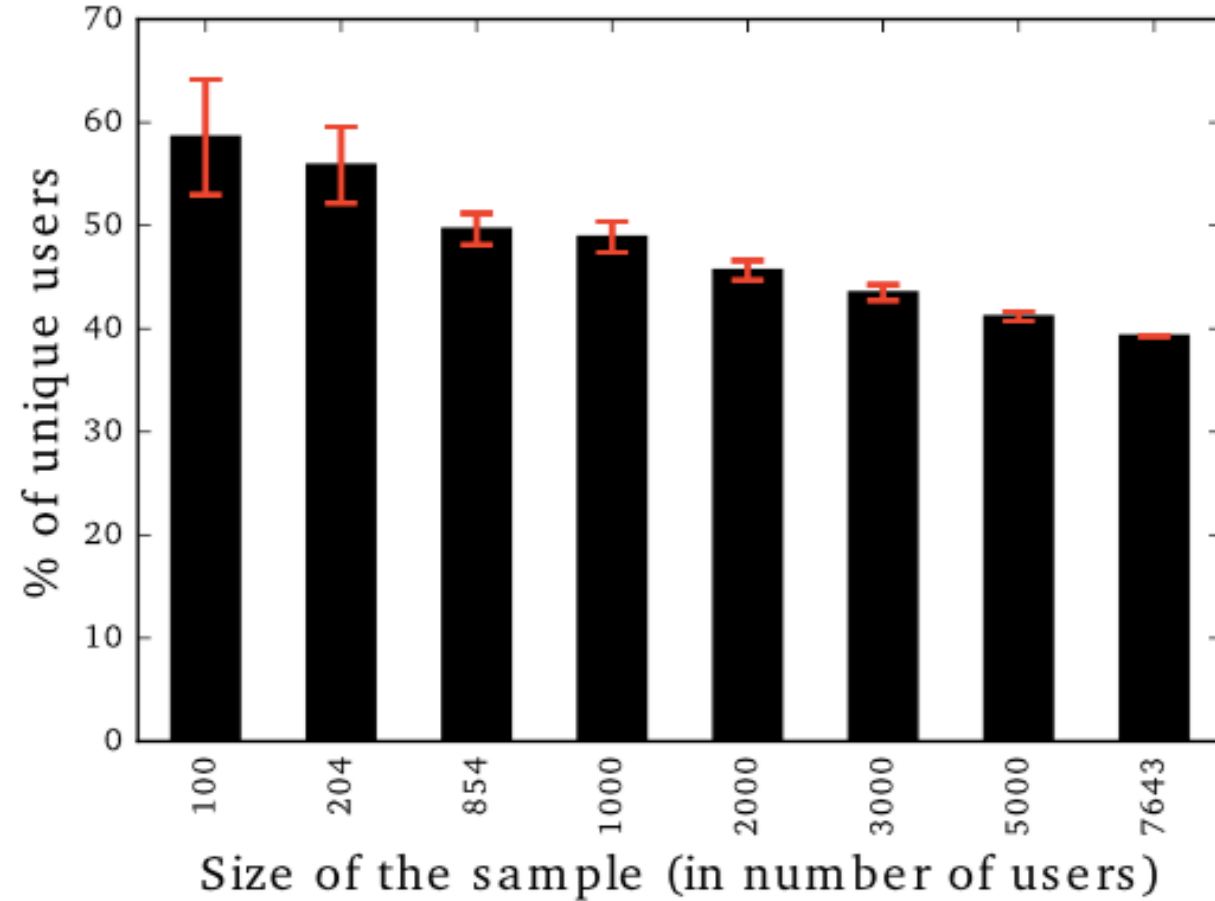
**Table 2: Previous studies on measuring uniqueness based on browser extensions and our estimation of uniqueness.**

Study	Fingerprints collected in a study	Extensions targeted in a study	Unique fingerprints in a study	Unique fingerprints in our dataset
Timing leaks [54]	204	2,000	56.86%	55.64%
XHOUND [58]	854	1,656	14.10%	49.60%
Ours	7,643	13k	39.29%	39.29%

Uniqueness grows  
as the dataset  
grows!

How to get a meaningful  
dataset?

How to define when we have  
enough users?

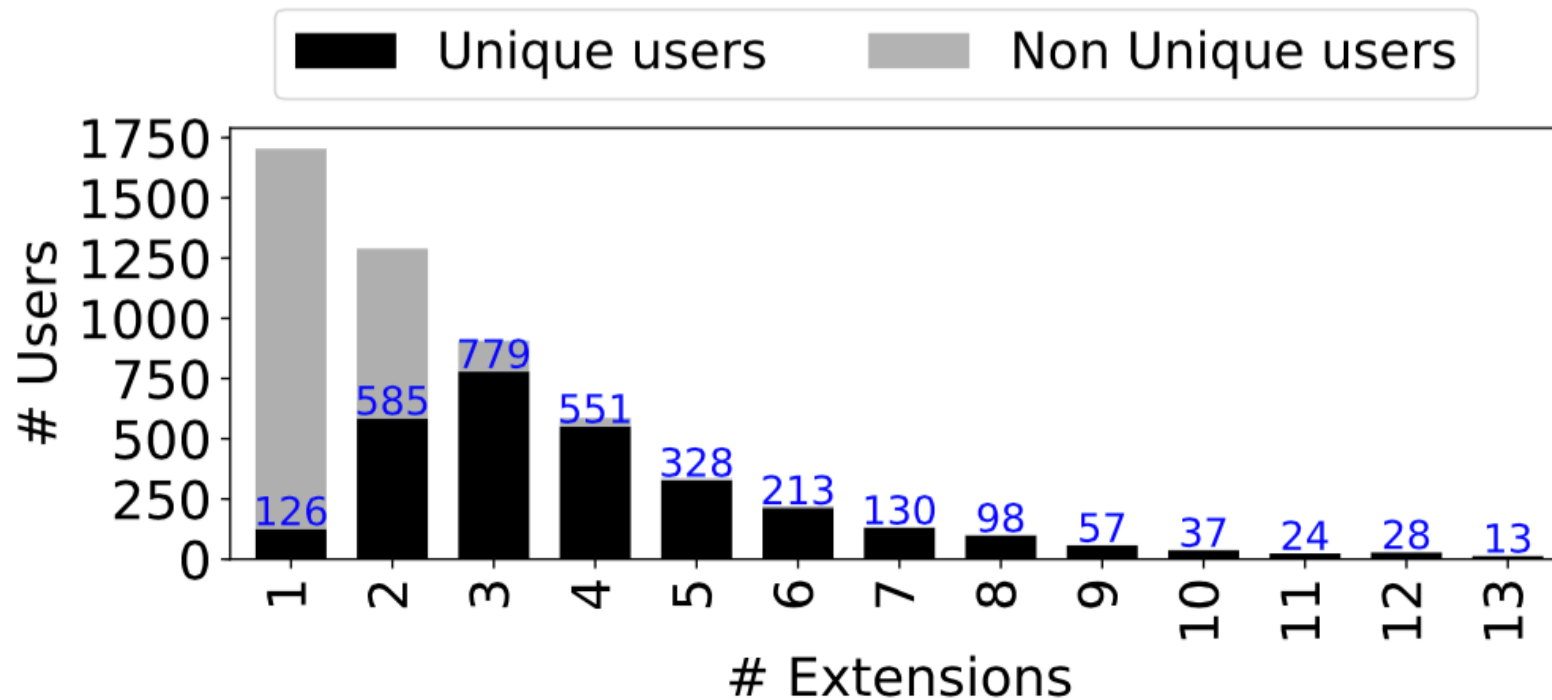


**Figure 13: Uniqueness of Chrome users based on their extensions only vs. number of users - 204 is the number of users used in [54] and 854 the number of users considered in [58]**

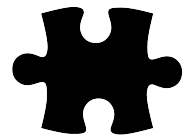
# How many extensions our users have?



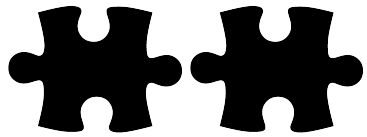
**7,643** users of Google Chrome browser



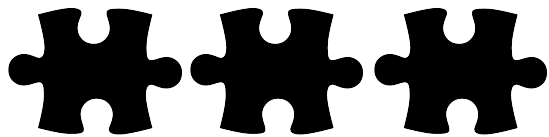
# Am I really unique if I use a few extensions?



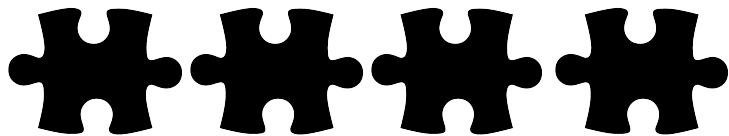
54.86% unique



76.25% unique



92.22% unique

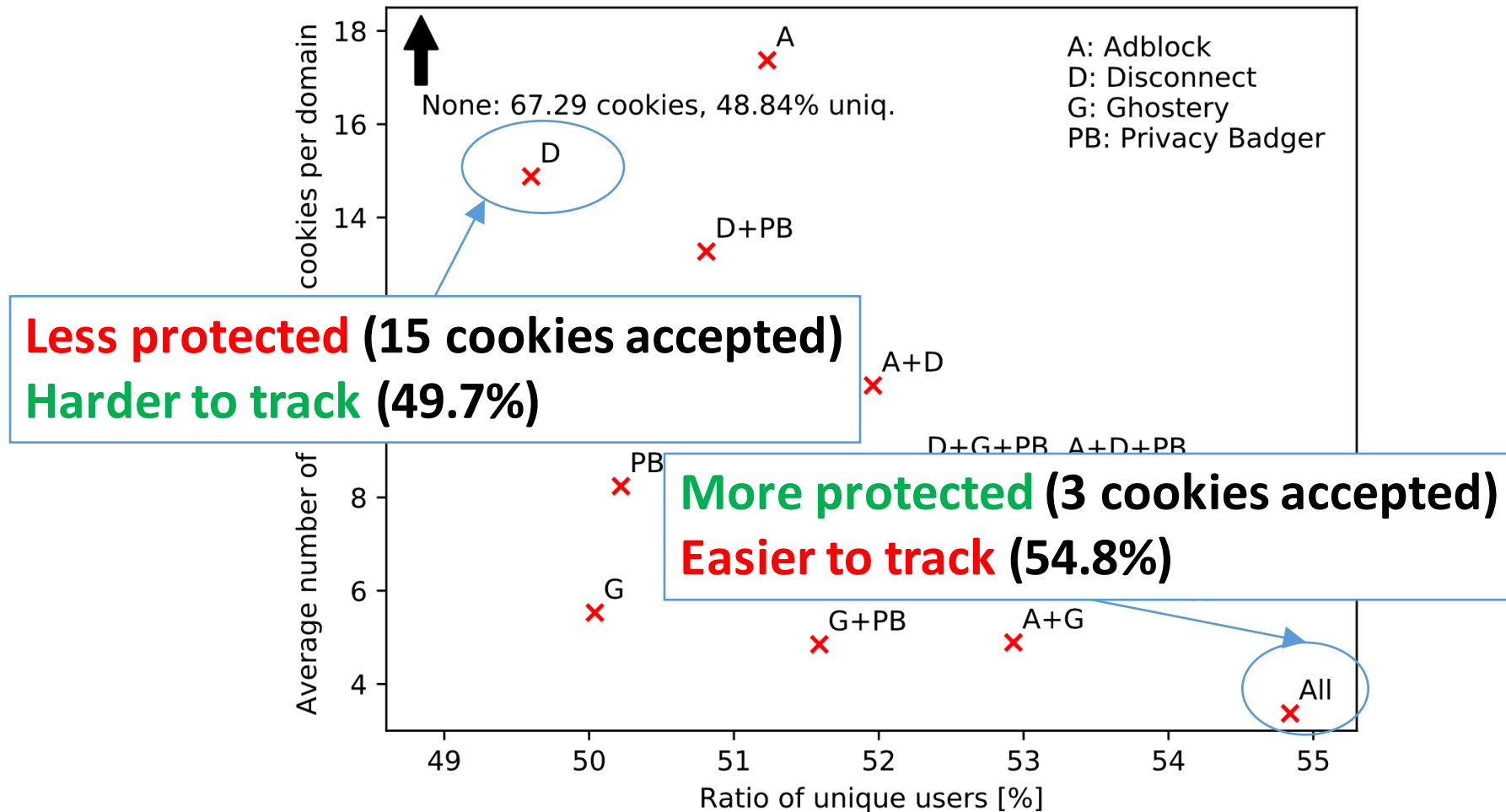


95.85% unique

# The dilemma of privacy extensions

- Privacy extensions **block some trackers**
- Privacy extensions **make a user more unique**
- What is the trade-off between **privacy gain** (some trackers are blocked) and **privacy loss** (user is more unique)?

# Uniqueness of users vs. number of accepted third-party cookies



\*4,000 pages crawled