WASHINGTON STATE UNIVERSITY

Peeking Through the Keyhole: Exploring Web-Based Privacy-Invasive Attacks

Xu Lin

Washington State University

CySER Seminar October 16, 2023

What security and privacy threats do we face everyday during our online activities?

Breaking and Fixing the Web

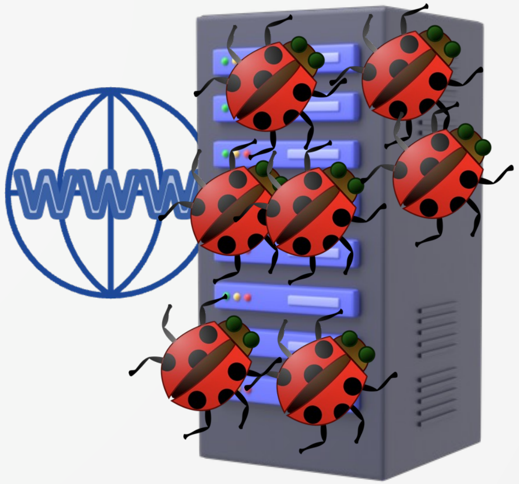
- Analyzing systems, mechanisms and protocols
  - Web browsers, web applications
  - Demonstrating novel and practical attacks against popular systems (incentivize major companies to better protect users)

- Developing privacy-preserving mechanisms & better security mechanisms

Privacy-invasive attacks

(high-level view)

- Browser features
- Browser fingerprints



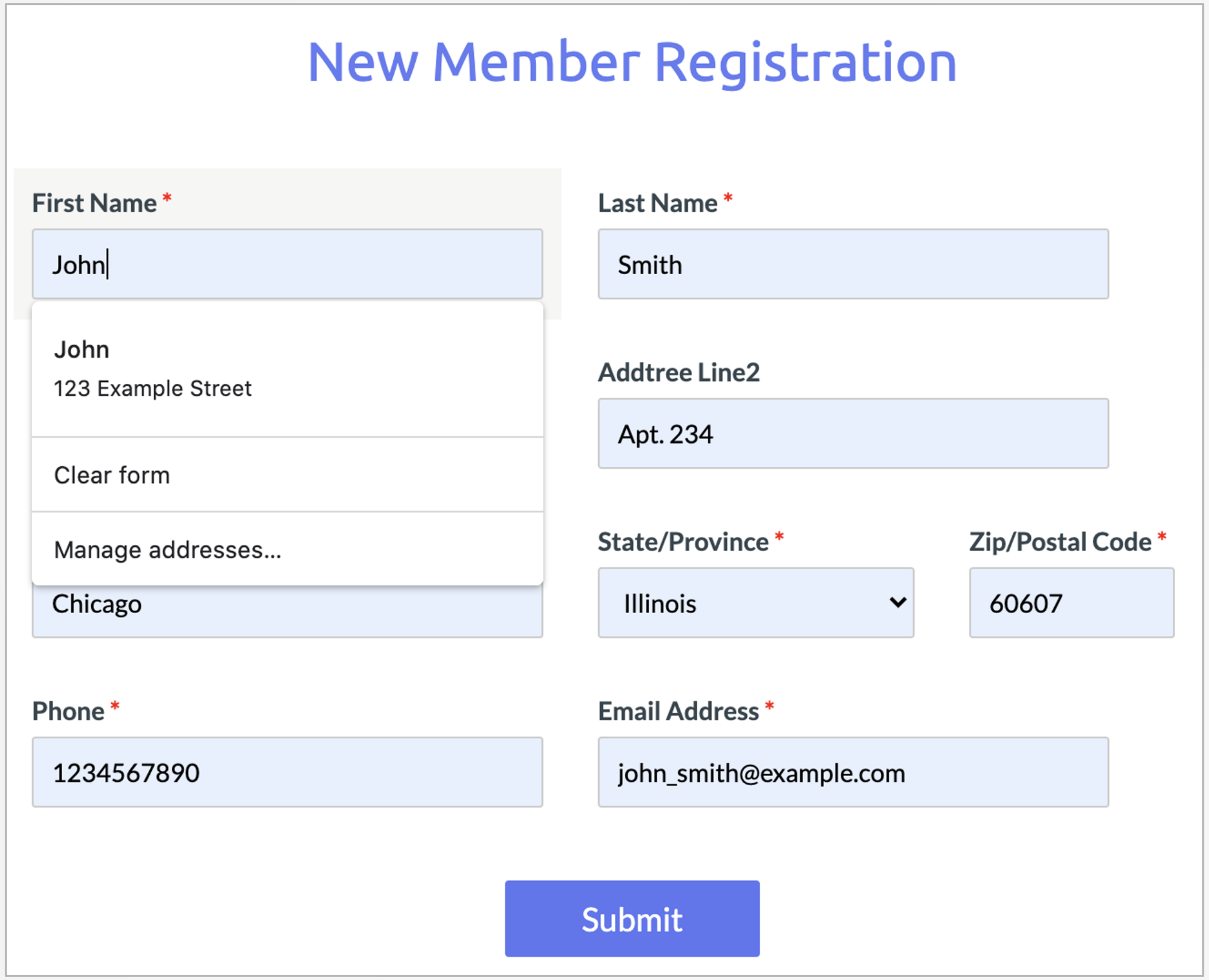
Time for a closer look…

Fill in the B l a n k s: Empirical Analysis of the Privacy Threats of Browser Form Autofill

Xu Lin, Panagiotis Ilia, Jason Polakis

ACM CCS 2020

Form autofill functionality



New Member Registration

First Name *
John

John
123 Example Street

Clear form

Manage addresses...

Chicago

Last Name *
Smith

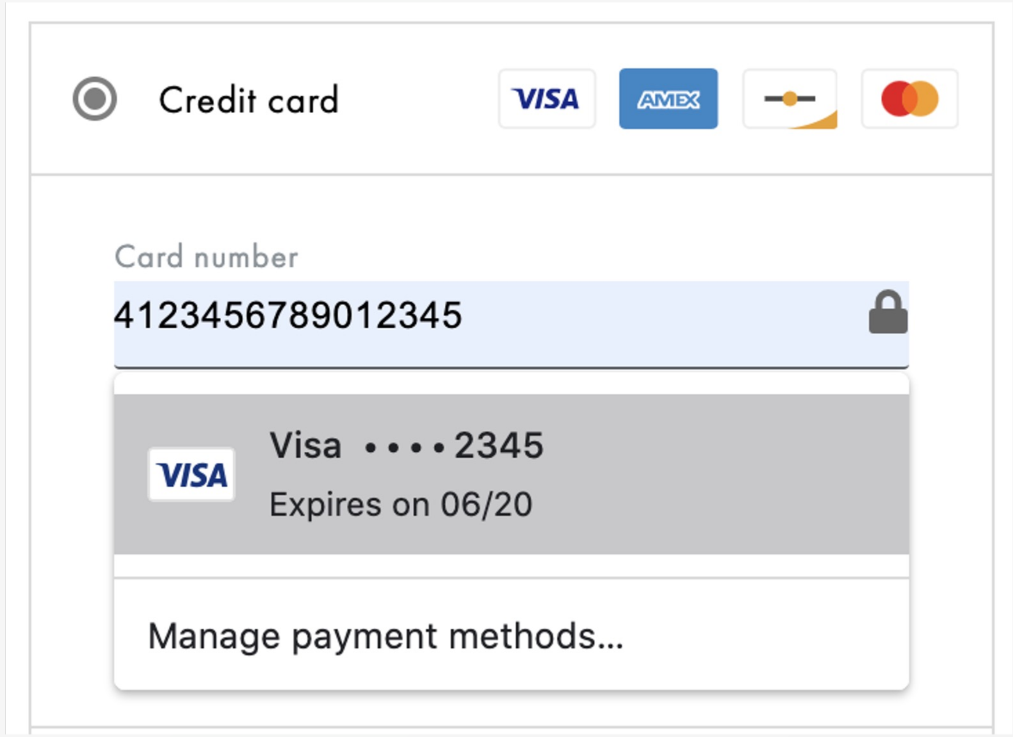
Addtree Line2
Apt. 234

State/Province *
Illinois

Zip/Postal Code *
60607

Phone *
1234567890

Email Address *
john_smith@example.com



◉ Credit card

Card number
4123456789012345

VISA Visa ••••2345
Expires on 06/20

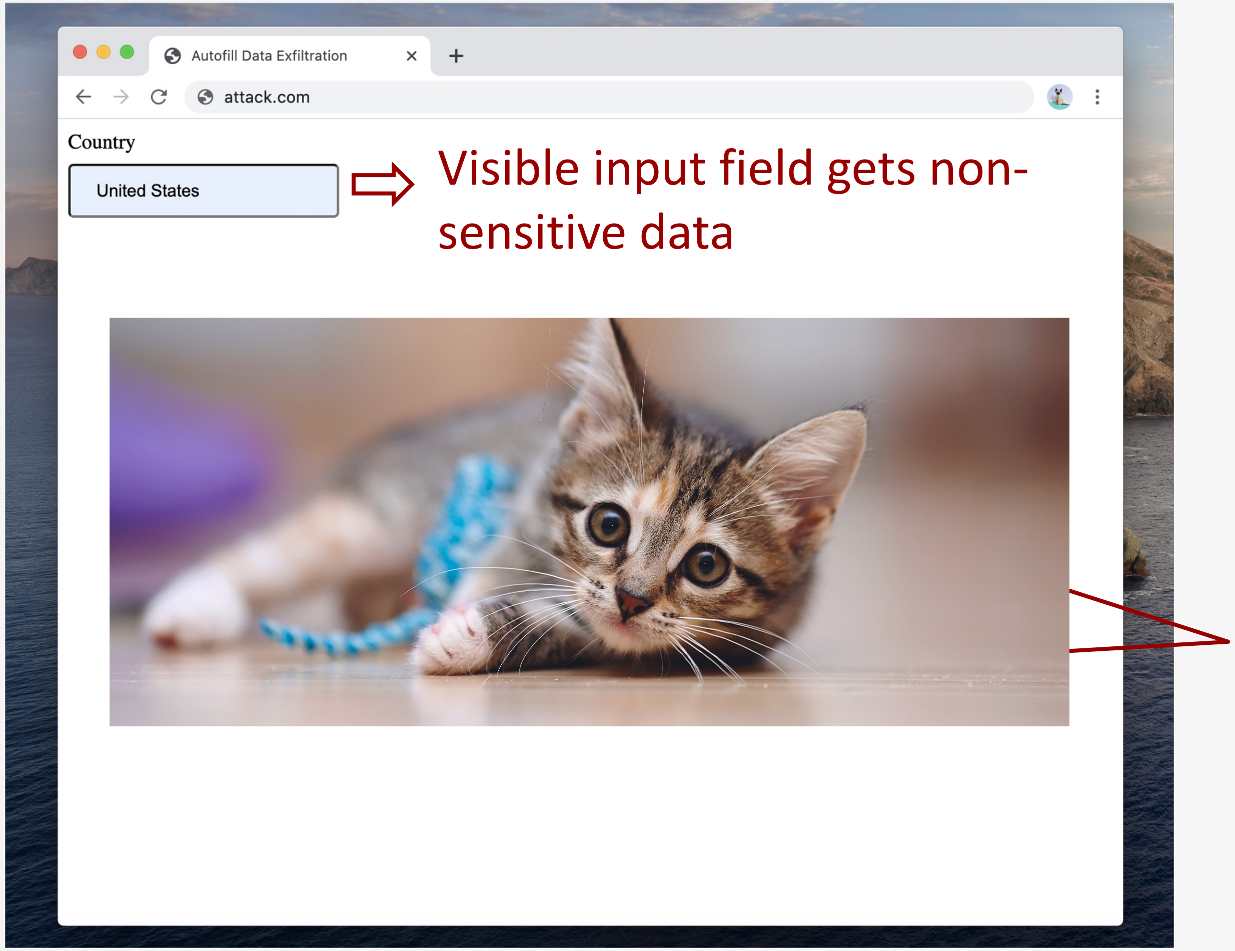
Manage payment methods...

Attack

- Malicious websites can obtain sensitive user data
    - Without the user’s **knowledge** or **consent**

- We demonstrate two types of attacks
    - ➢ Stealthy data exfiltration
    - ➢ Data inference attack

Stealthy data exfiltration: **visually hidden** form elements



Country

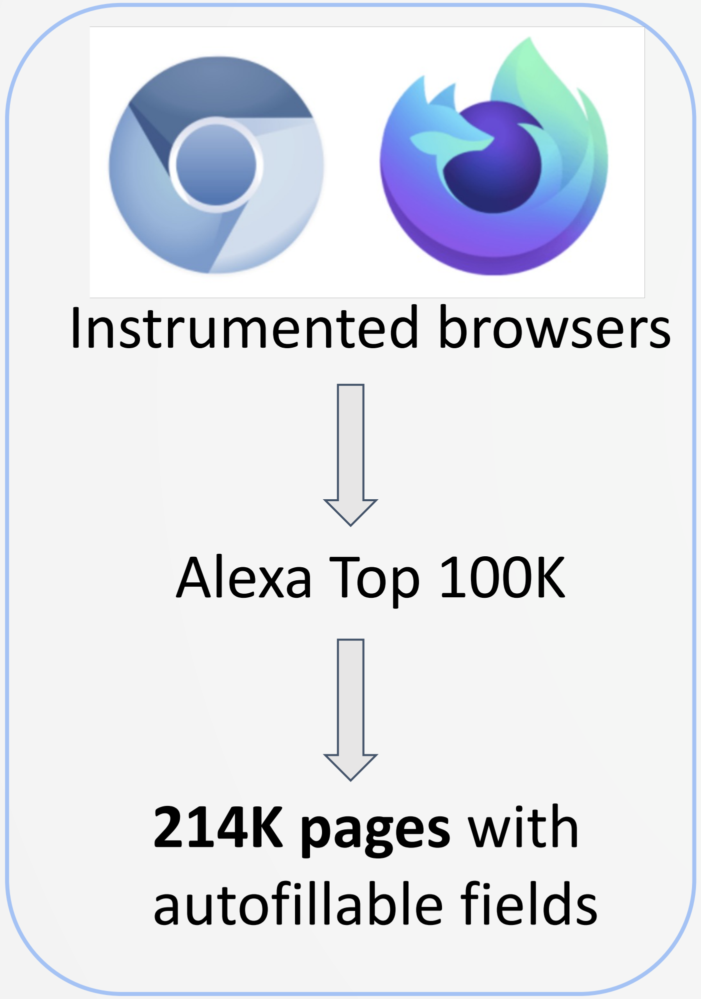
United States

Visible input field gets non-sensitive data

- Hidden form elements in the page
- Filled automatically by browsers when autofill is triggered

Hidden input fields get sensitive data

Visually hidden elements - Measurement

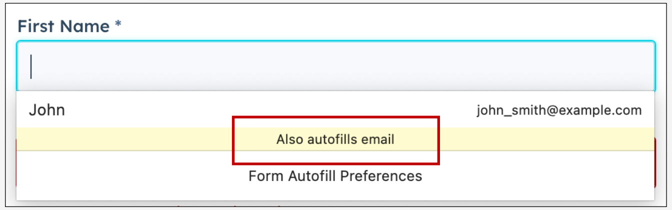

Instrumented browsers

Alexa Top 100K

214K pages with autofillable fields

	Firefox	Chrome
Sites w/ autofilled forms	21,589	31,621
Sites w/ hidden fields	**24.52%**	**5.82%**

- **Chrome** fills forms in **46.5%** more websites
  - ➢ does not respect *autocomplete="off"*
- **Firefox** fills almost **3x** forms with **hidden fields**
  - ➢ no visibility check
  - ➢ displays a message

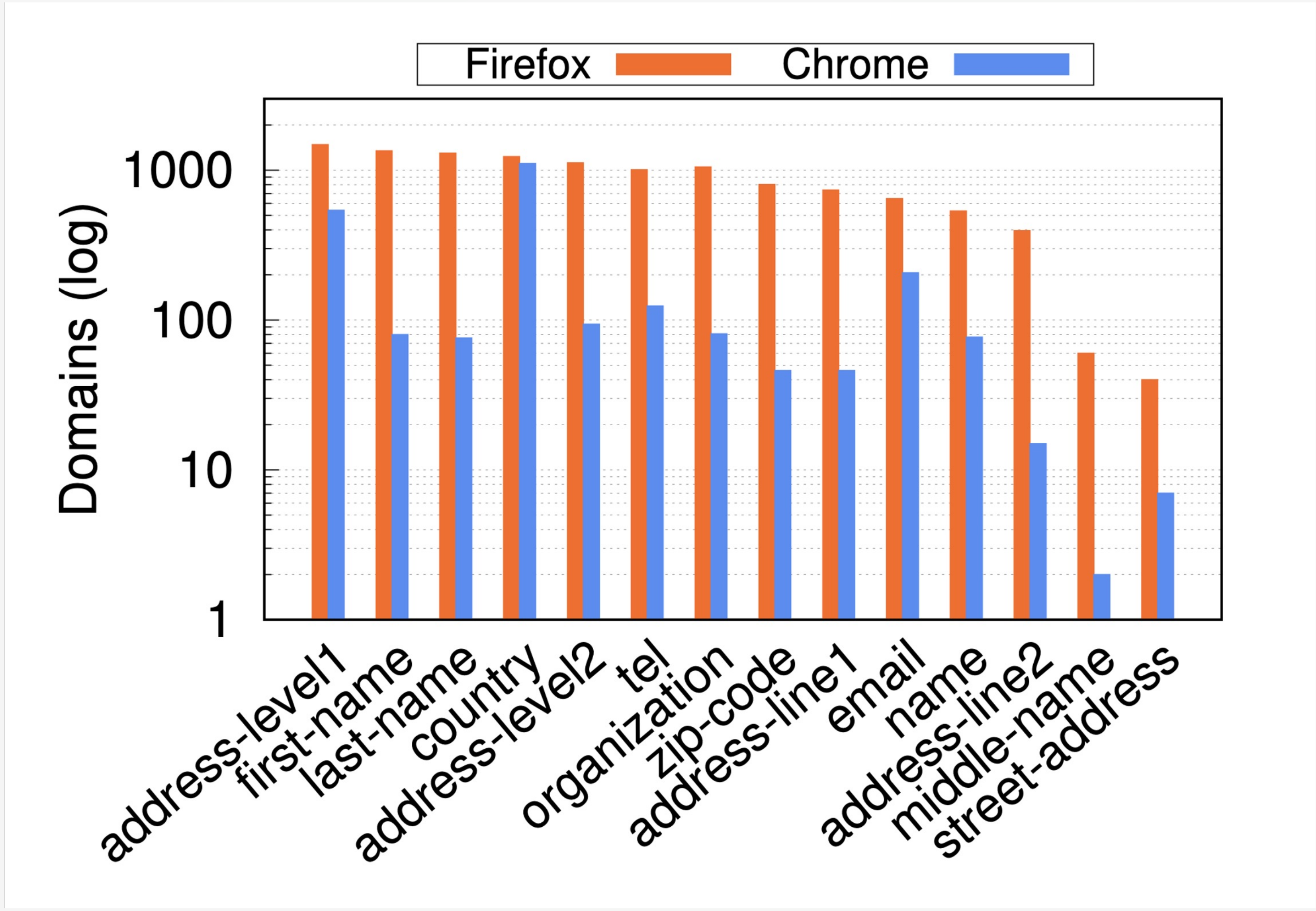

First Name *

John john_smith@example.com

Also autofills email

Form Autofill Preferences

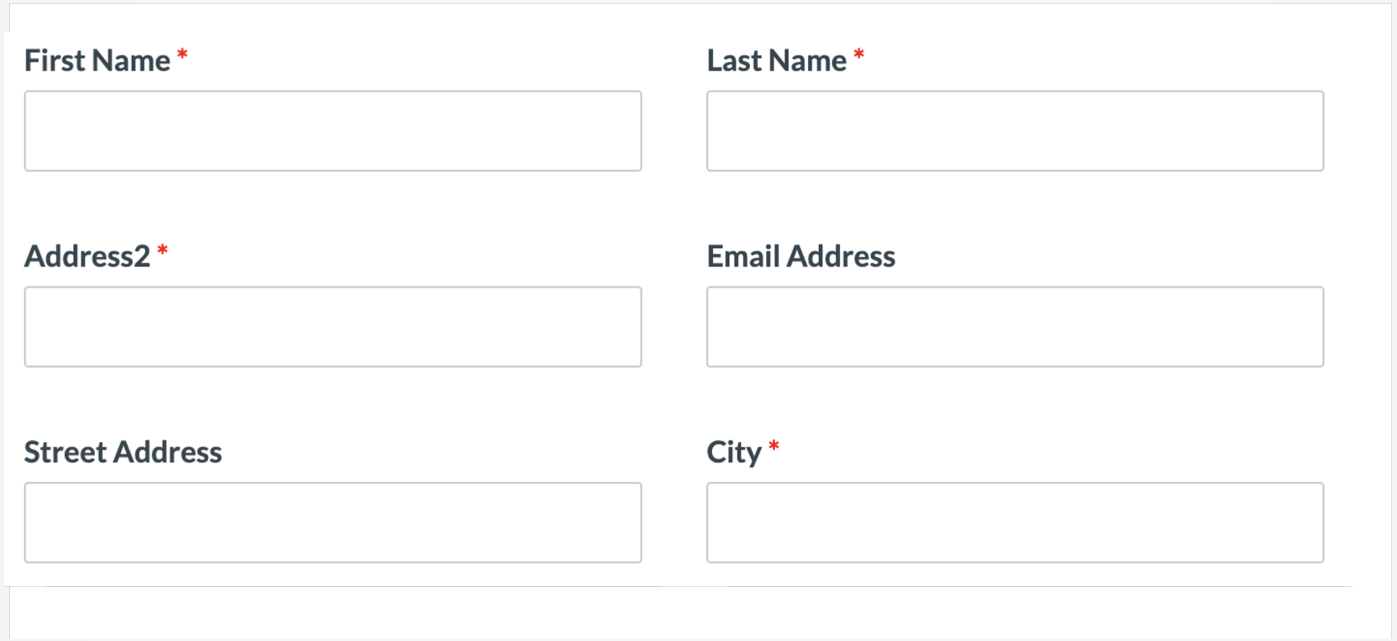
Types of hidden autofilled fields



Cautious users may avoid using autofill.

Are they safe?

Autofill preview feature



First Name *

Last Name *

Address2 *

Email Address

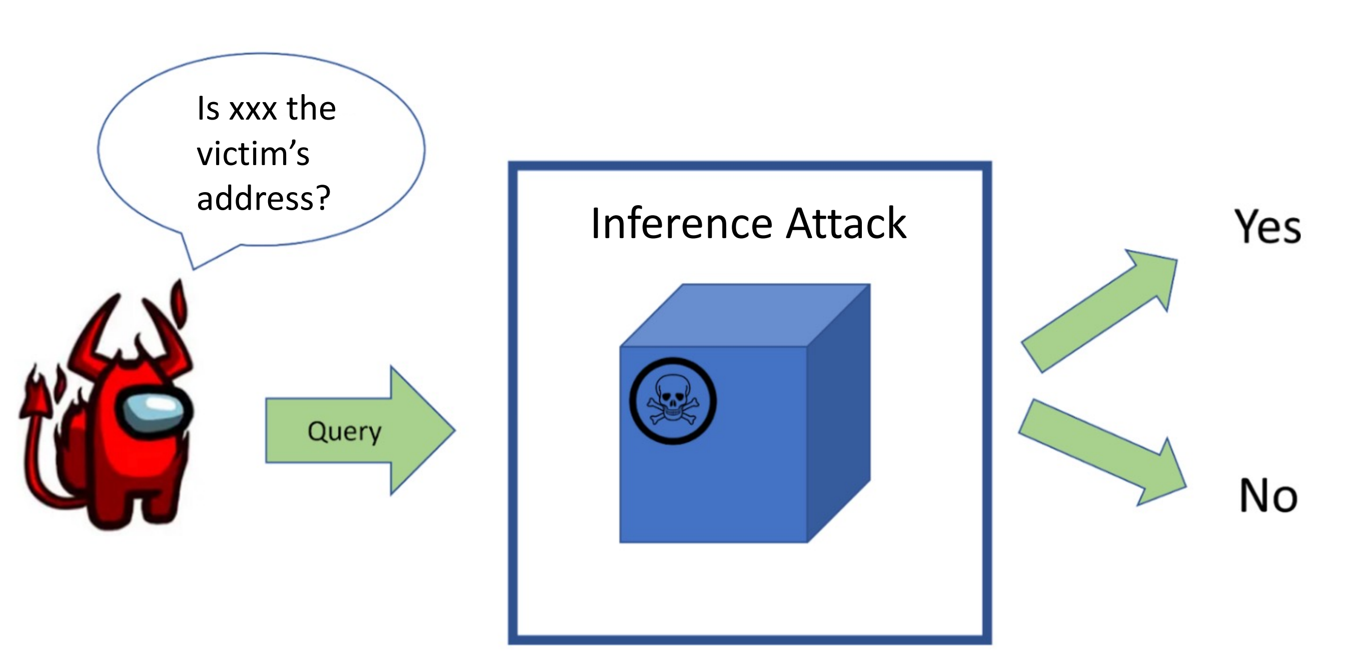
Street Address

City *

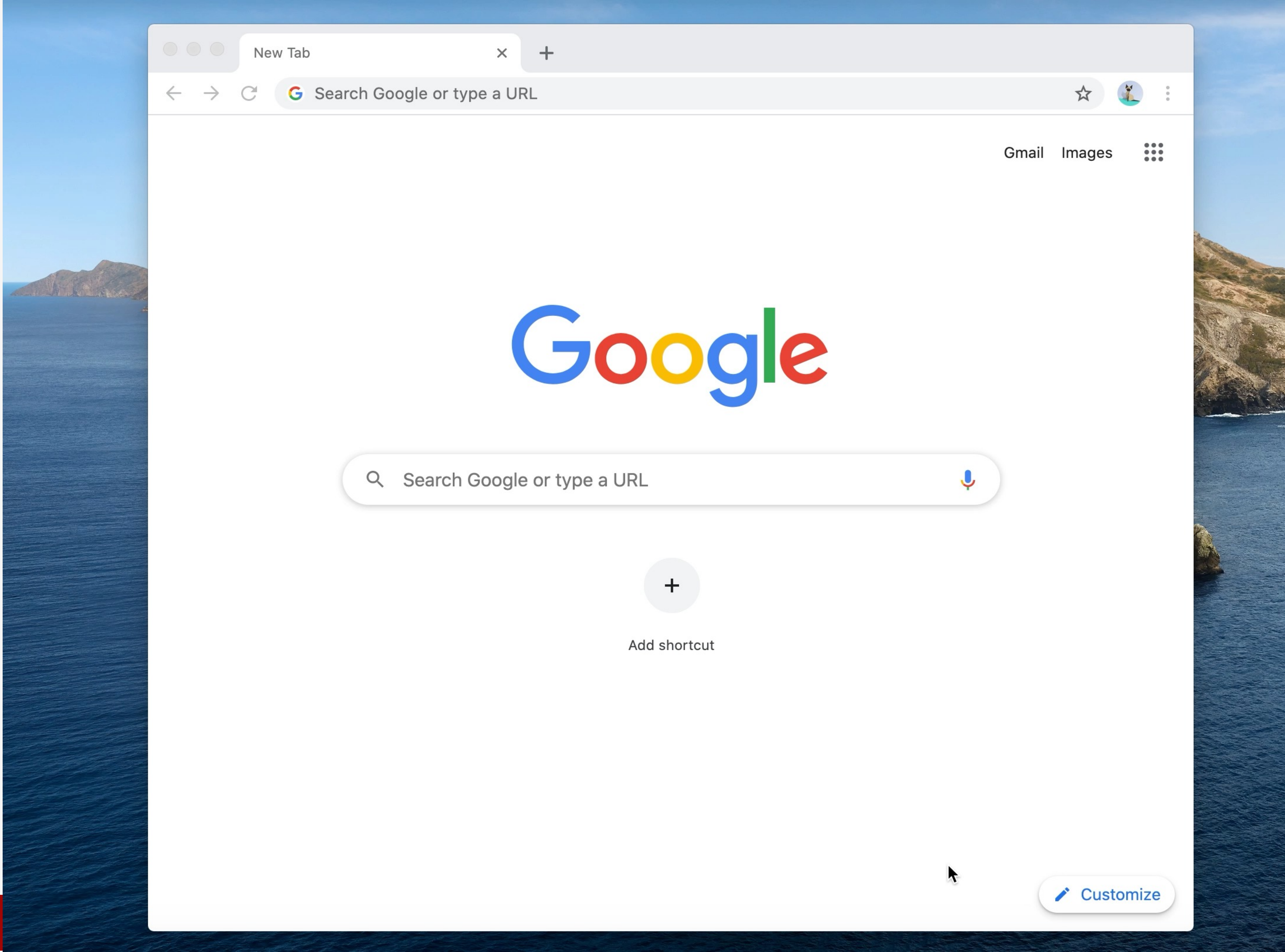
- Preview values are displayed in overlay fields that are **not** part of the DOM
  - **Not** accessible to the page (i.e., through JavaScript)

Data inference attack exploits **autofill preview** functionality

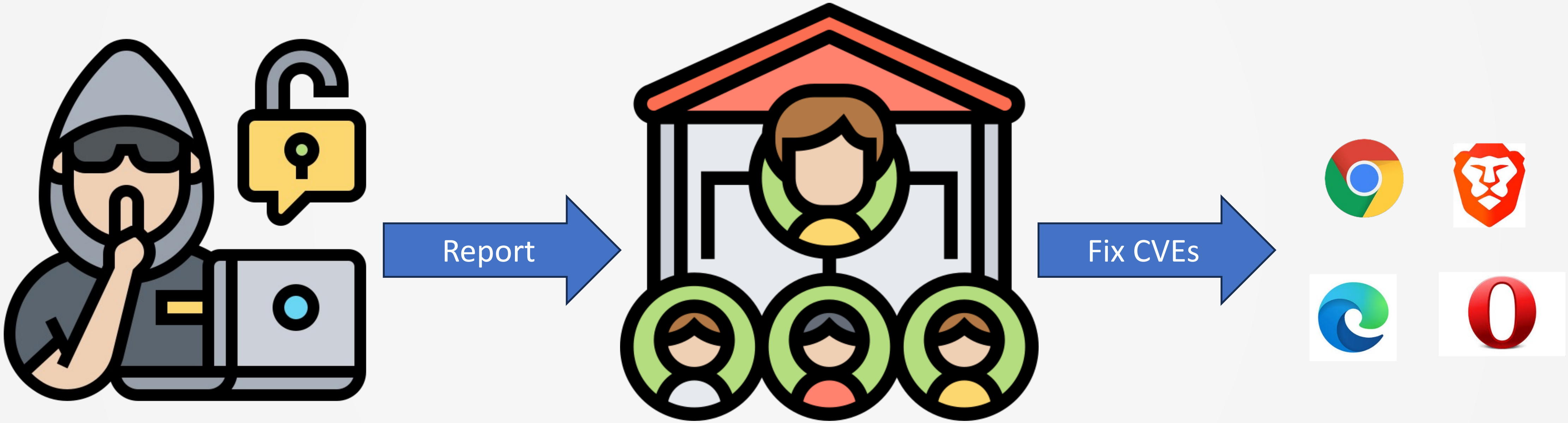
- Does not require users to trigger autofill
- Runs when user clicks on a field and values are previewed



Data Inference Attack



Breaking and Fixing the Web



Report

Fix CVEs

Phish in Sheep's Clothing: Exploring the Authentication Pitfalls of Browser Fingerprinting

Xu Lin, Panagiotis Ilia, Saumya Solanki, Jason Polakis

USENIX Security 2022

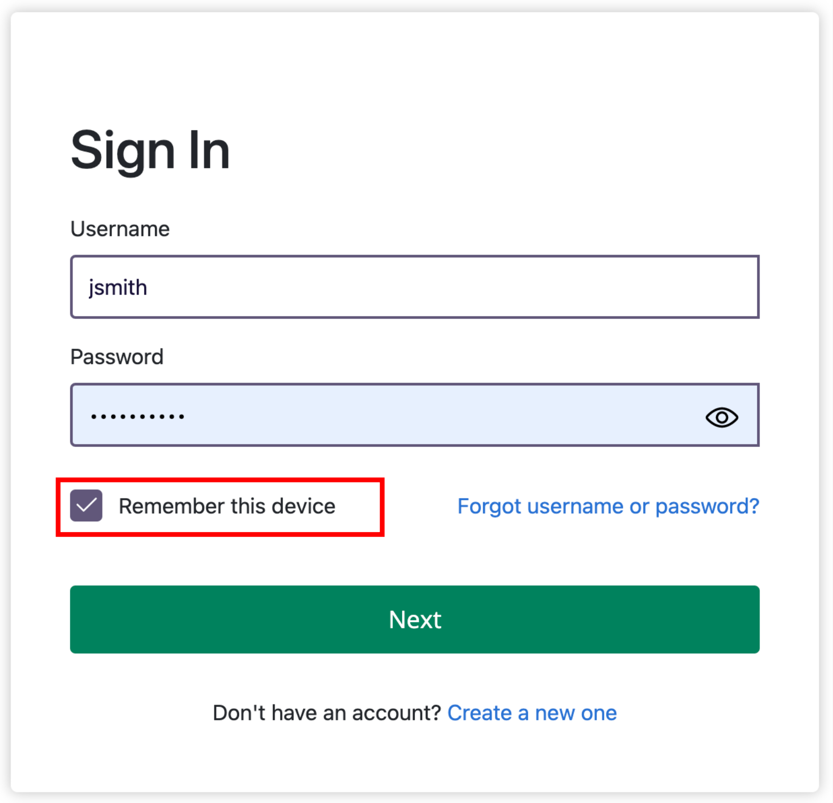
image: www.welzie-art.com/

Motivation

- Account hijacking remains a major problem
- Phishing is a prevalent hijacking vector [1,2]
- Two-factor authentication (2FA) is a *critical* defense
  - Device-based challenges block >94% of phishing-based hijacking attempts, 100% of automated hijacking attempts [3]

[1] Bursztein et al. "Handcrafted fraud and extortion: Manual account hijacking in the wild." *IMC '14*.
[2] Thomas et al. "Data breaches, phishing, or malware? understanding the risks of stolen credentials." *CCS '17*
[3] Doerfler et al., "Evaluating login challenges as a defense against account takeover. " *WWW '19*

Risk-based authentication and two-factor authentication (2FA)



Sign In

Username

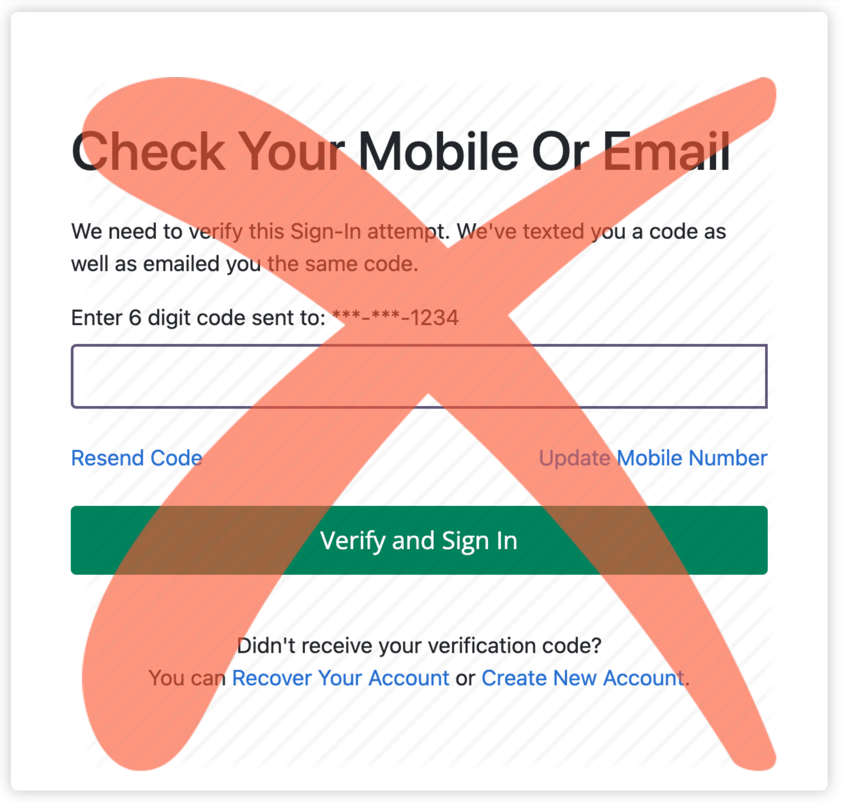
jsmith

Password

••••••••••

☑ Remember this device

Don't have an account? Create a new one



Check Your Mobile Or Email

We need to verify this Sign-In attempt. We've texted you a code as well as emailed you the same code.

Enter 6 digit code sent to: ***-***-1234

Didn't receive your verification code?
You can Recover Your Account or Create New Account.

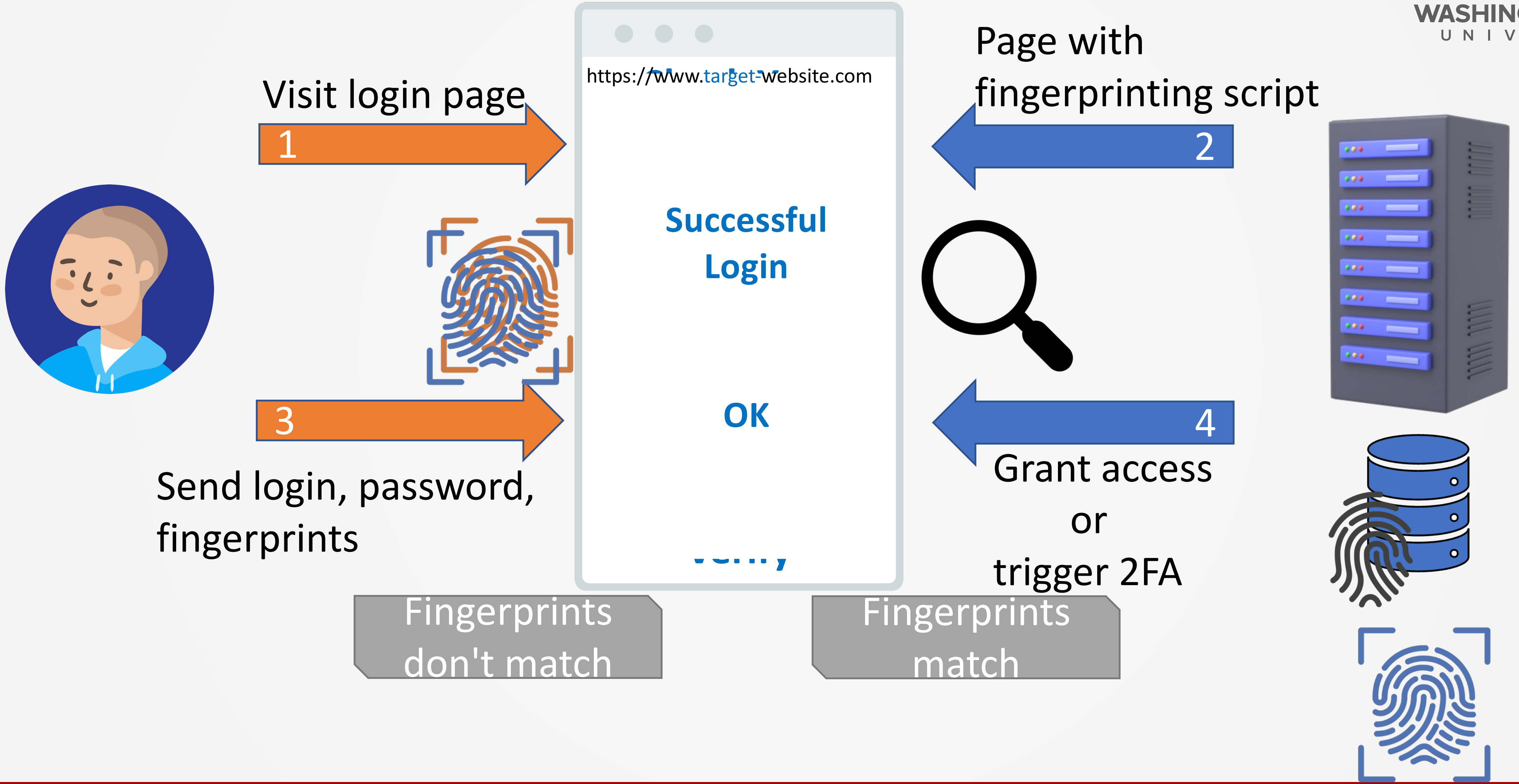
- 2FA creates friction for users
- Certain websites only trigger 2FA for *suspicious* login attempts

Browser fingerprints

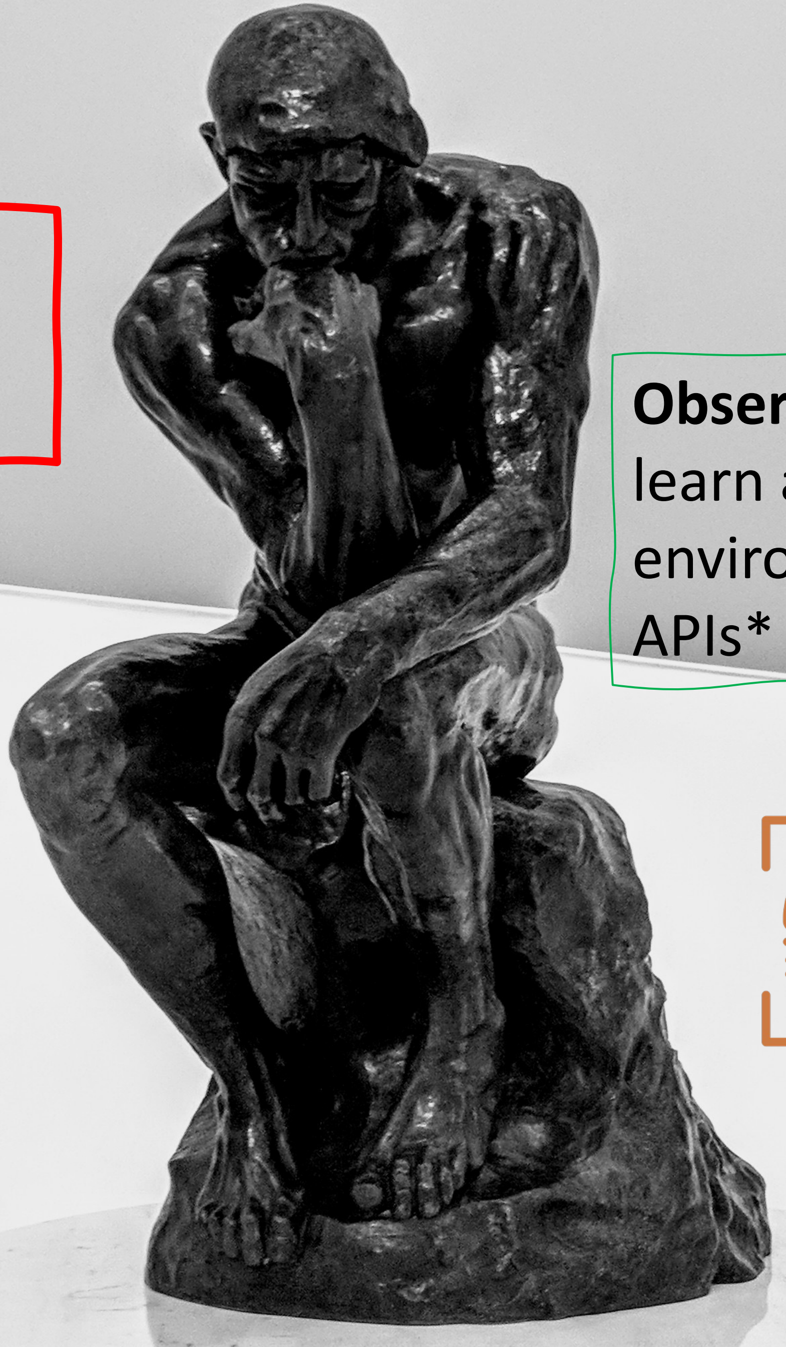
“A device fingerprint, machine fingerprint, or browser fingerprint is information collected about a remote computing device for the purpose of identification. Fingerprints can be used to fully or partially ***identify individual users or devices*** *even when cookies are turned off.”*

- Browser fingerprints can be trivially collected by ***any*** website the user visits through a series of **JavaScript APIs**.

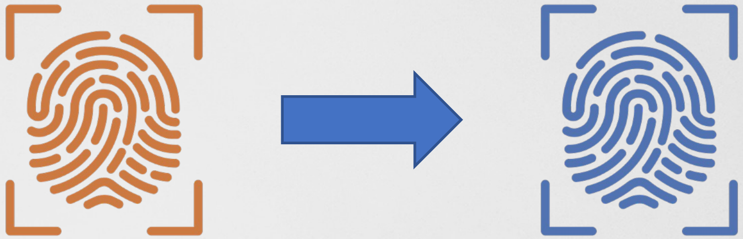
Advanced risk-based authentication that uses browser fingerprinting



What can attackers do to trick websites into **not** considering a login suspicious?



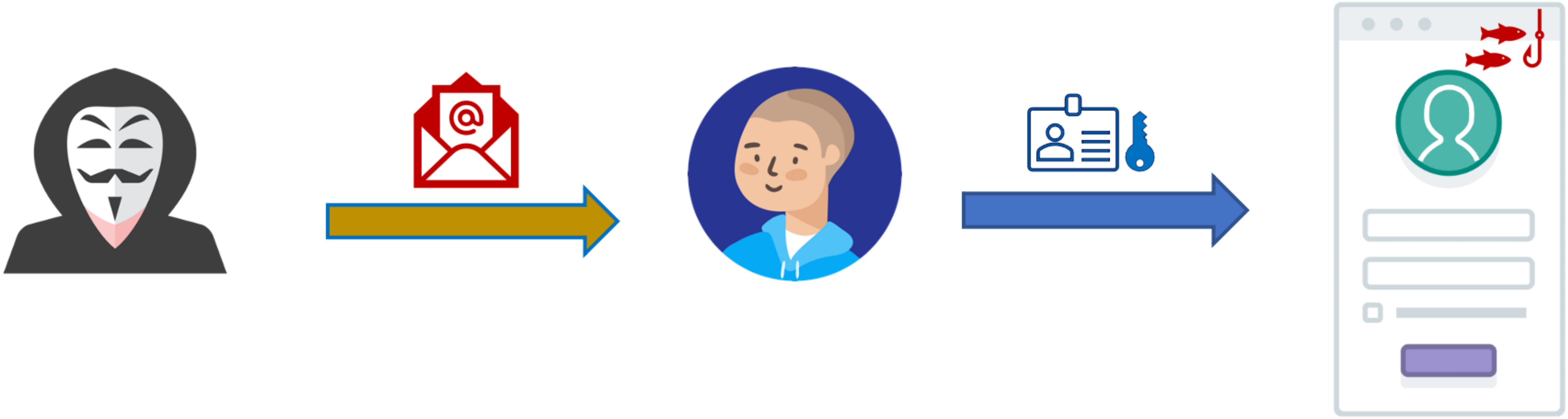
Observation: websites can only learn about the user's environment through browser APIs* available to any website.



* And HTTP headers, which can be easily spoofed.

Threat Model

The attacker tricks the user into visiting a malicious website and entering their credentials.



Overview of our attack workflow

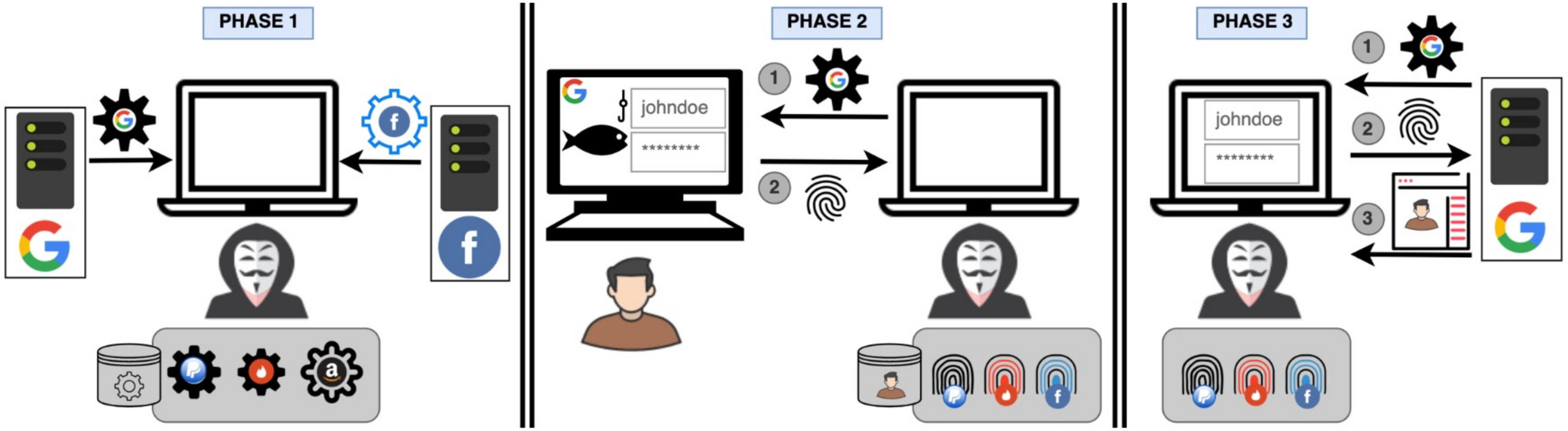
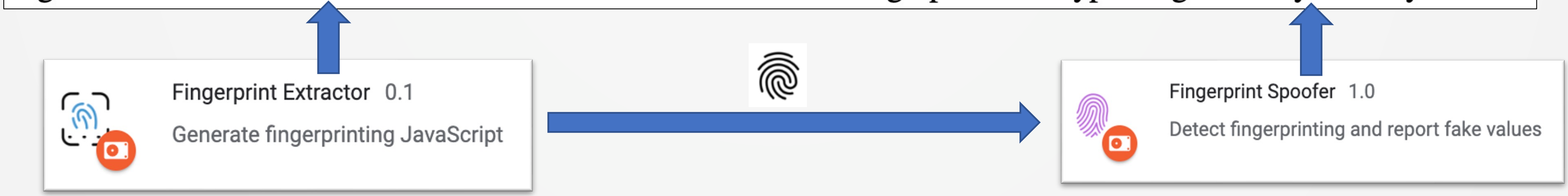


Figure 1: Overview of our attack workflow that misuses browser fingerprints for bypassing ancillary security checks.



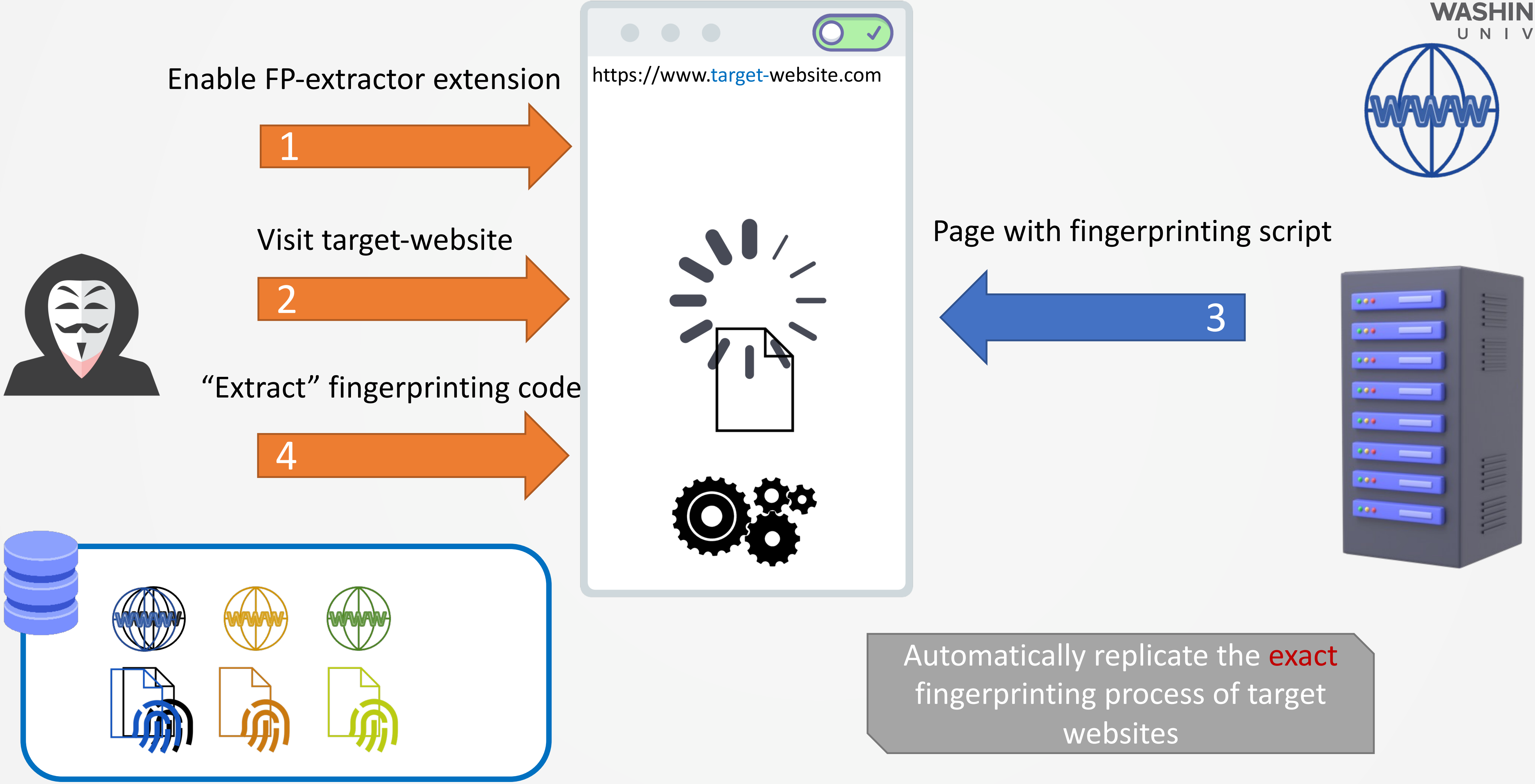
Fingerprint Extractor 0.1

Generate fingerprinting JavaScript

Fingerprint Spoofer 1.0

Detect fingerprinting and report fake values

Phase1: attacker visits target websites and "extracts" their fingerprinting code



Enable FP-extractor extension

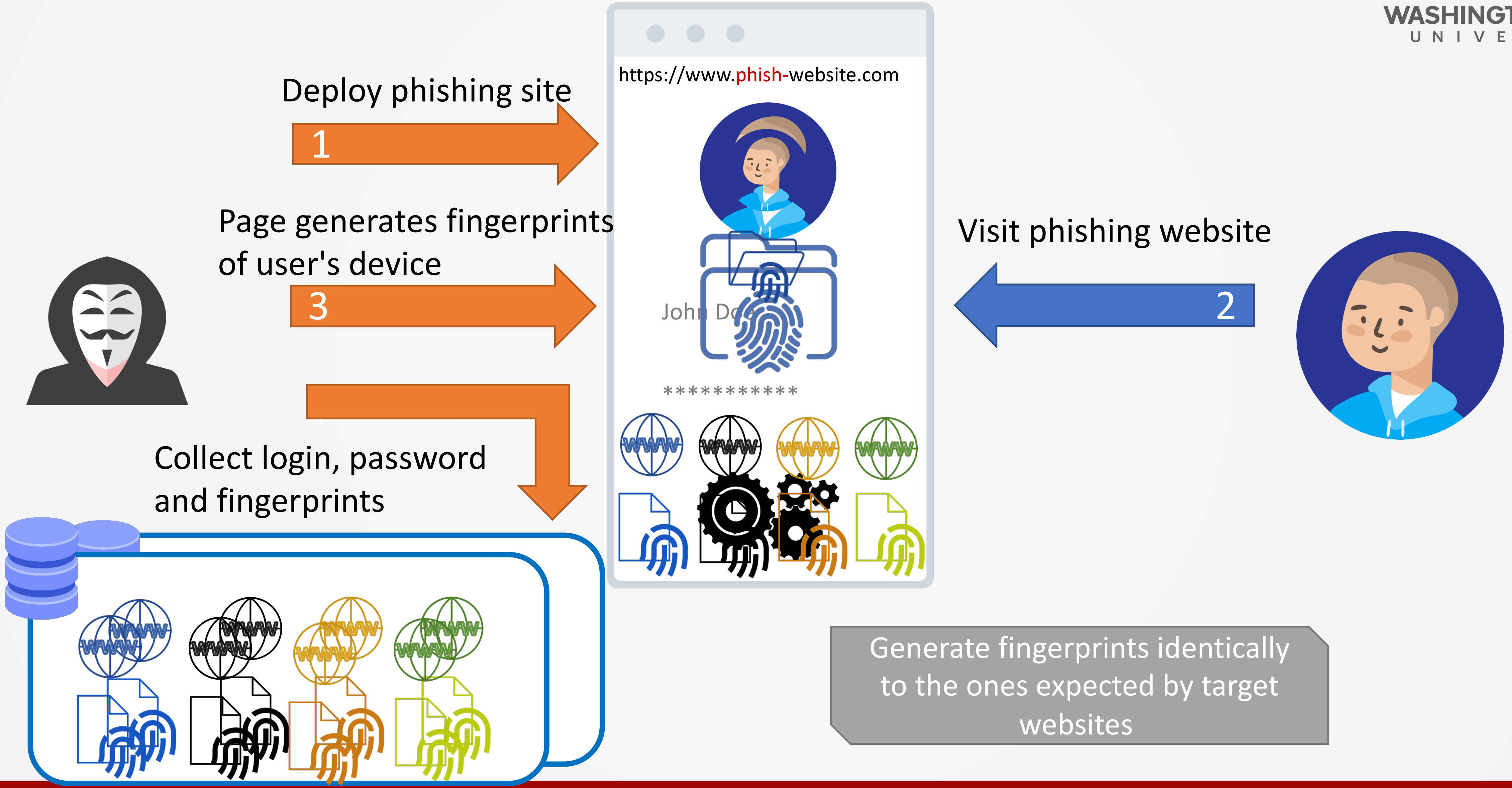
Visit target-website

Page with fingerprinting script

“Extract” fingerprinting code

Automatically replicate the exact fingerprinting process of target websites

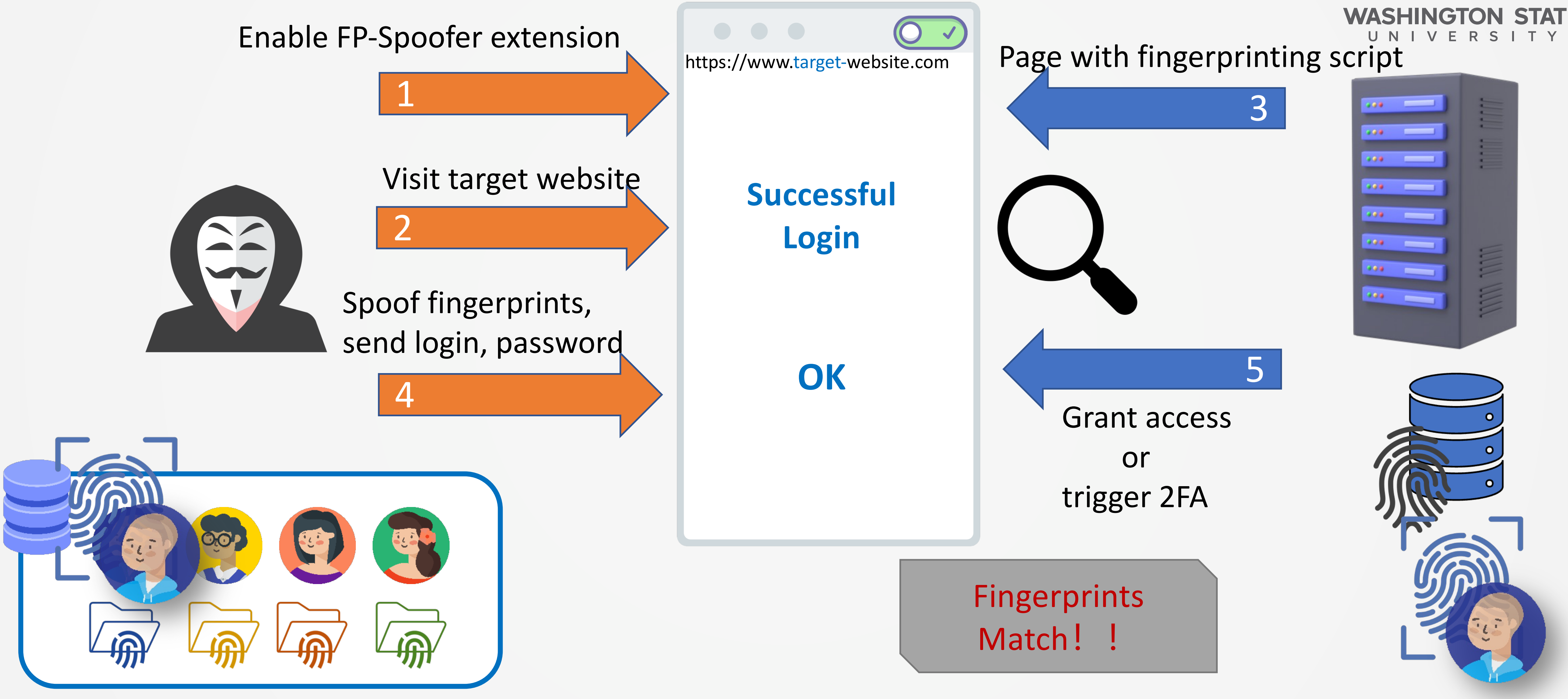
Phase2: attacker obtains user’s credentials and fingerprints



Deploy phishing site
1
https://www.phish-website.com
Page generates fingerprints of user's device
3
John Doe
****
Visit phishing website
2
Collect login, password and fingerprints
www

Generate fingerprints identically to the ones expected by target websites

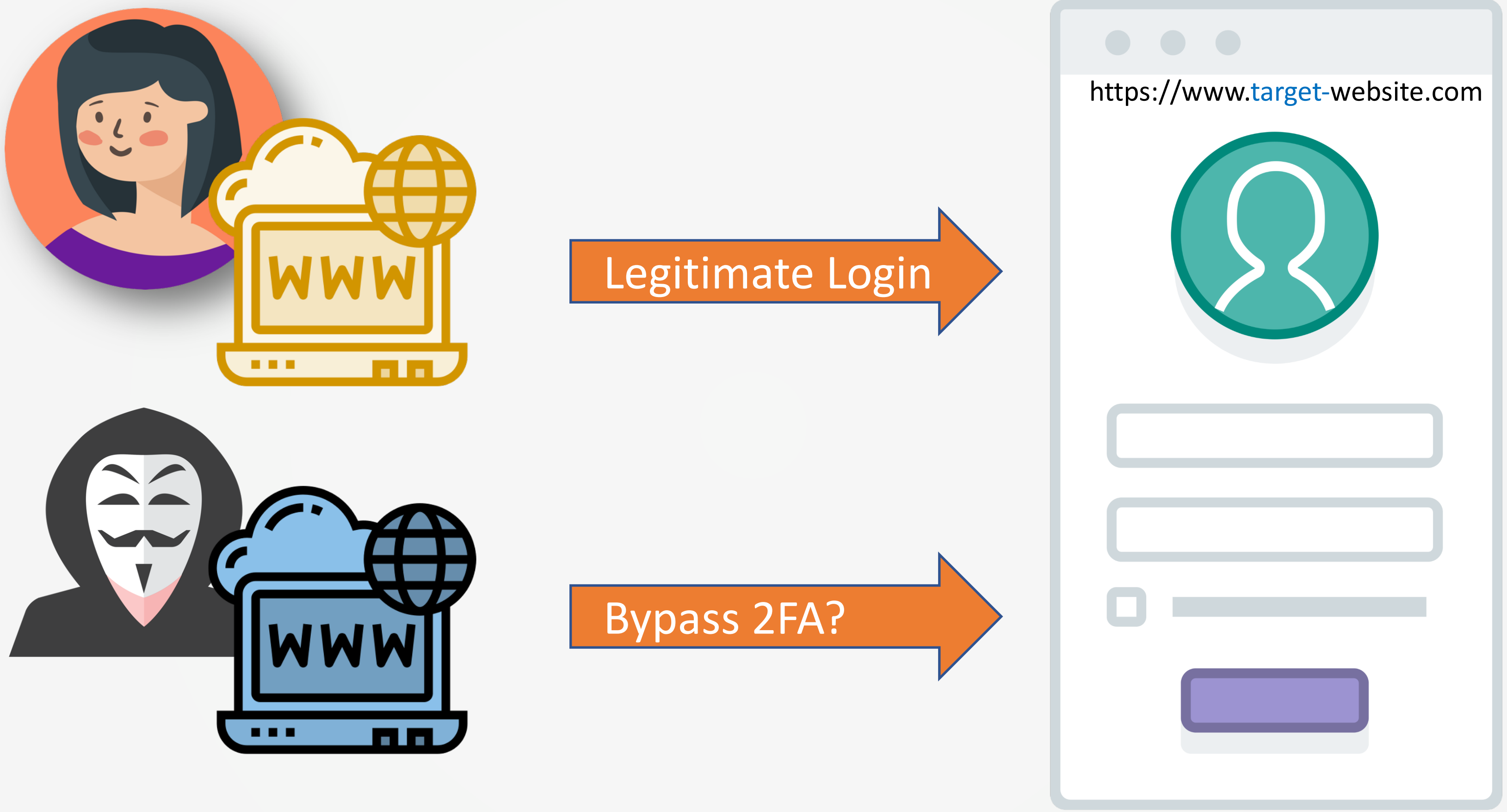
Phase3: attacker spoofs fingerprints and bypasses 2FA mechanism



Fingerprint Spoofing Demo

attacker spoofs their device’s fingerprints to mimic those of the victim’s device

What about the real world ?



Legitimate Login

Bypass 2FA?

https://www.target-website.com

Risk-based authentication mechanisms in popular web services

Website	Fingerprinting Technique				IP Address Restrictions		Vulnerable
	BasicFP	Canvas/WebGL	Fonts	Audio	IP Check	Bypass	
Bank-A	✓	✗	✗	✗	✗	-	✓
Bank-B	✗	✗	✗	✗	✓	✗	✗
CreditCard	✓	✗	✗	✗	✓	→	✓
Trading-A	✓	✗	✗	✗	✗	-	✓
Trading-B	✗	✗	✗	✗	✓	→	✓
Tax-A	✓	✓	✗	✗	✓	✗	✗
Tax-B	✓	✓	✓	✗	✗	-	✓
Tax-C	✓	✓	✓	✓	✗	-	✓
Tax-D	✓	✓	✓	✓	✓	✗	✗
eCommerce-A	✓	✓	✗	✗	✗	-	✓
eCommerce-B	✓	✗	✗	✗	✓	✗	✗
RideSharing	✓	✓	✓	✗	✓	→	✓
Food&Beverage-A	✓	✗	✗	✗	✓	○	✓
Food&Beverage-B	✓	✗	✗	✗	✓	✗	✗

- **We completely bypass 2FA in 9/14 websites that use FPs for authentication!**
- Attack only prevented by IP address checks.
- We inject X-Forwarded-For header (used by proxies) with the user’s IP to bypass IP-checks (→).
- Certain sites only require an IP from the same city (○).

What about phishing sites in the wild?

Phishing and Fingerprinting

Dataset	Time Period	Sites	JS	FP
Phish-A	31/05/2018 – 19/06/2019	71,343	39,618	29,312
Phish-B	31/10/2018 – 05/05/2020	82,431	40,777	36,733
APWG	05/05/2020 – 12/04/2021	173,269	93,568	85,491

- The majority collect fingerprints, with **73.98%**, **90.08%** and **91.36%** across the 3 datasets respectively.
- An **increase** in the number of websites collecting browser fingerprints over time.

Phishing sites that obtain all necessary fingerprints for bypassing 2FA

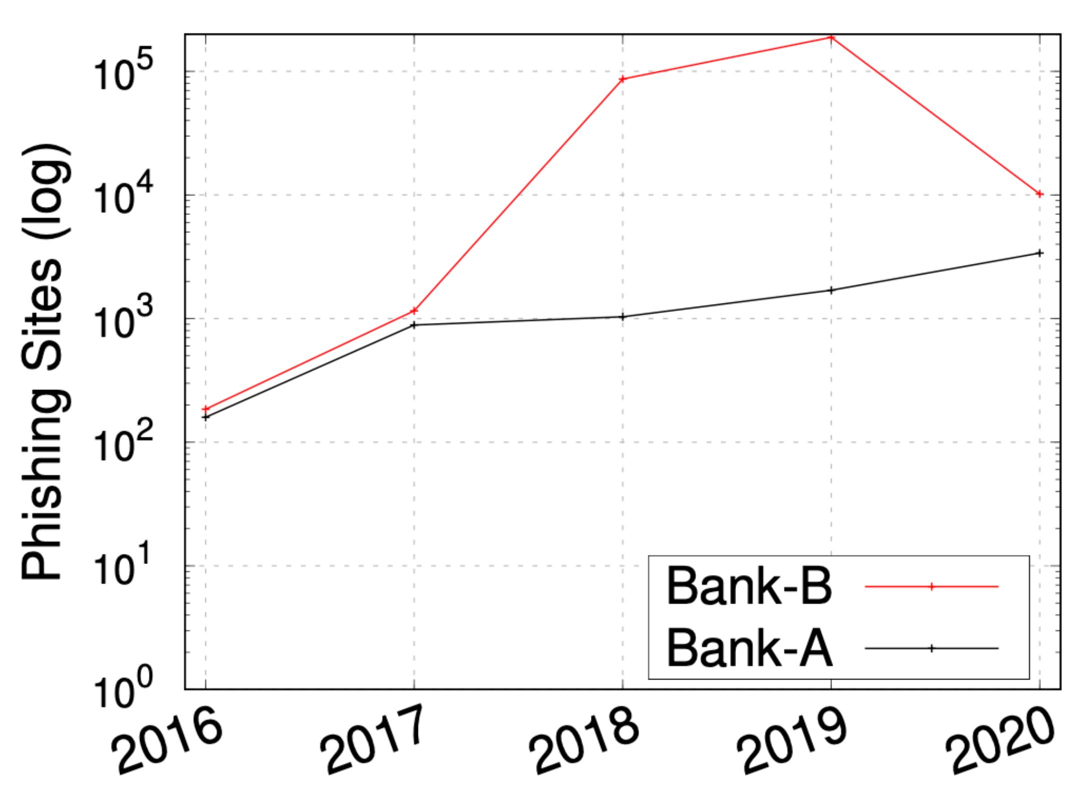
	Phish-A		Phish-B		APWG	
Target	Sites	Bypass	Sites	Bypass	Sites	Bypass
Bank-A	83	1	685	14	330	74
Bank-B	1549	-	2,683	-	327	-
CreditCard	89	61	0	0	12	0
Trading-A	0	0	0	0	6	6
RideSharing	7	0	363	1*	1378	5*
WebInfrastructure	0	0	1	1	220	219

APWG dataset

- more recent
- visited actual websites

* Indicates a mismatch in the arguments passed to fingerprinting functions.

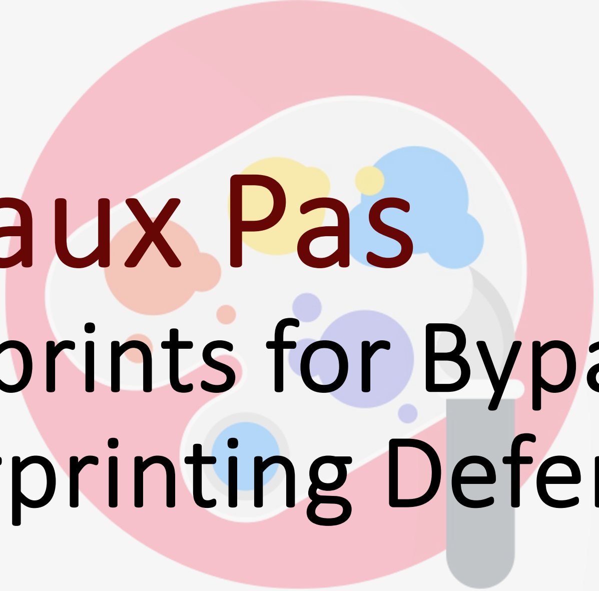
Are phishers adapting their targets?



- Bank-A is vulnerable to our attack.
- The sharp decline in phishing sites targeting Bank-B could be due to the IP address requirement.

Fixing the Web

- *Disclosure* of findings to affected vendors



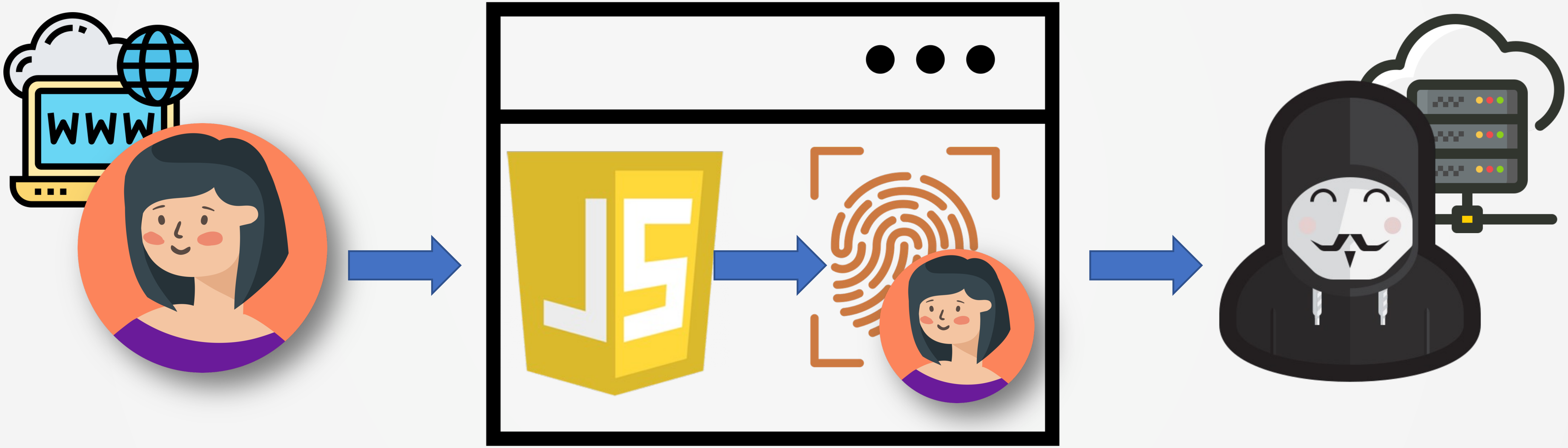
Fashion Faux Pas

Implicit Stylistic Fingerprints for Bypassing Browsers’ Anti-Fingerprinting Defenses

Xu Lin*, Frederico Araujo † , Teryl Taylor †, Jiyong Jang † , Jason Polakis*

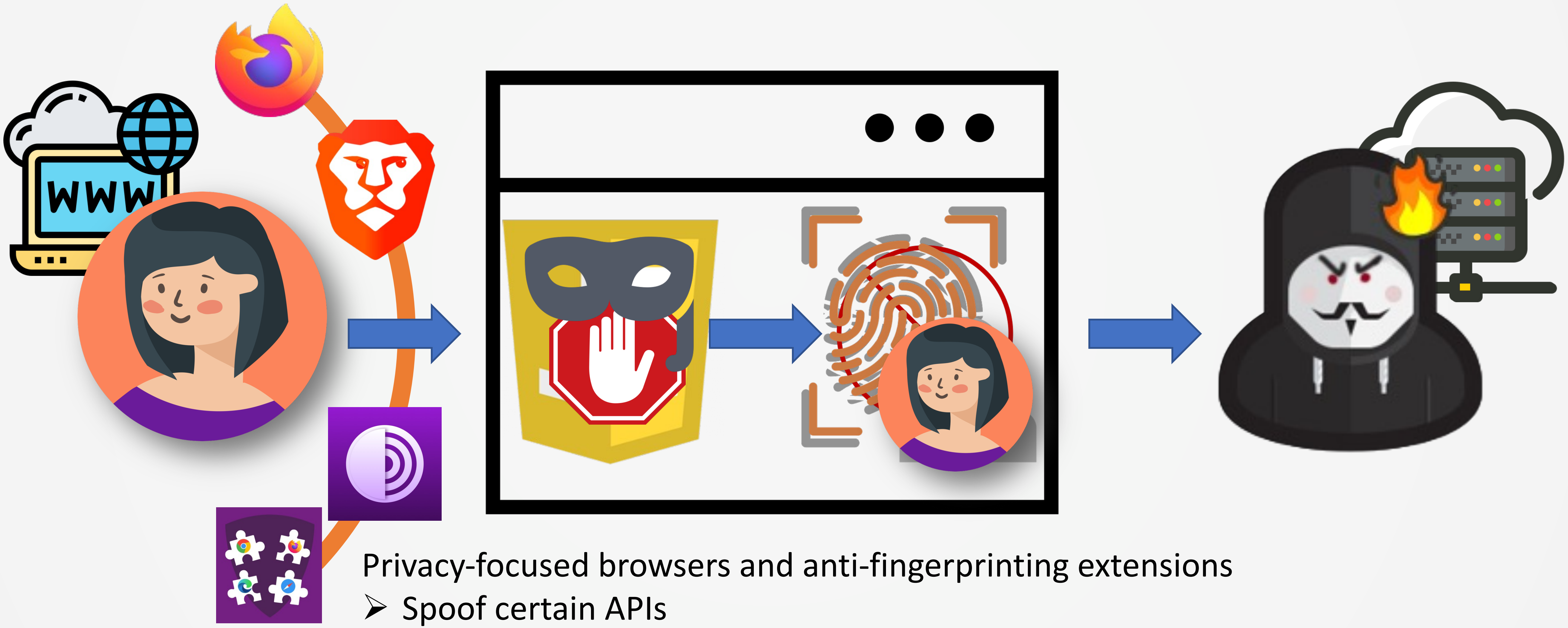
IEEE Symposium on Security and Privacy 2023

Online tracking



Browser fingerprinting heavily relies on JavaScript.

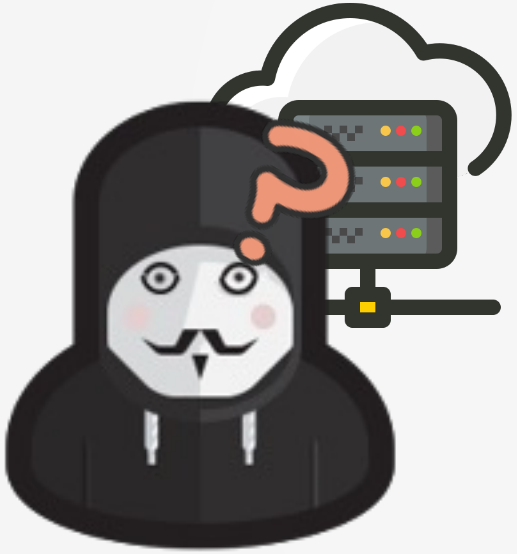
Fingerprinting countermeasures

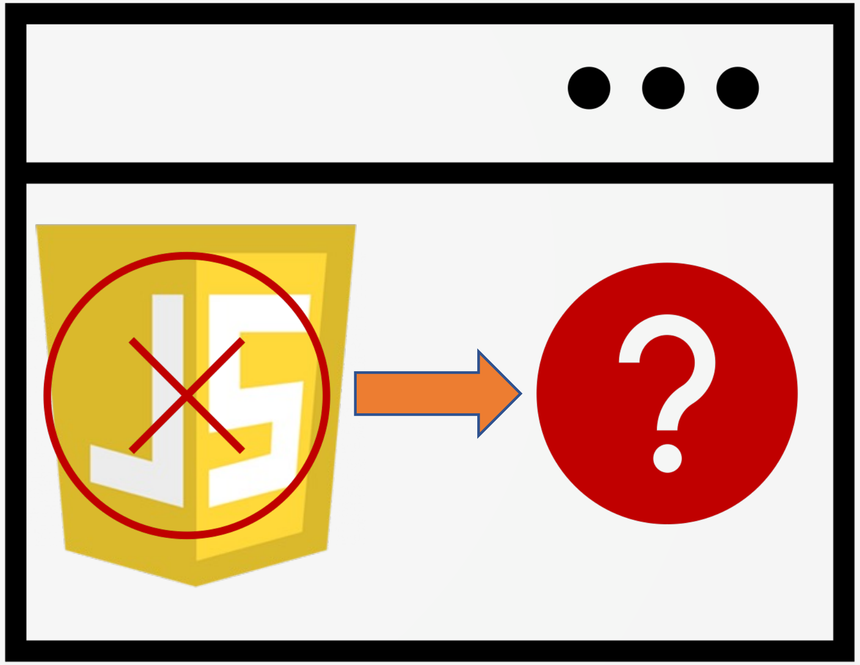


Privacy-focused browsers and anti-fingerprinting extensions

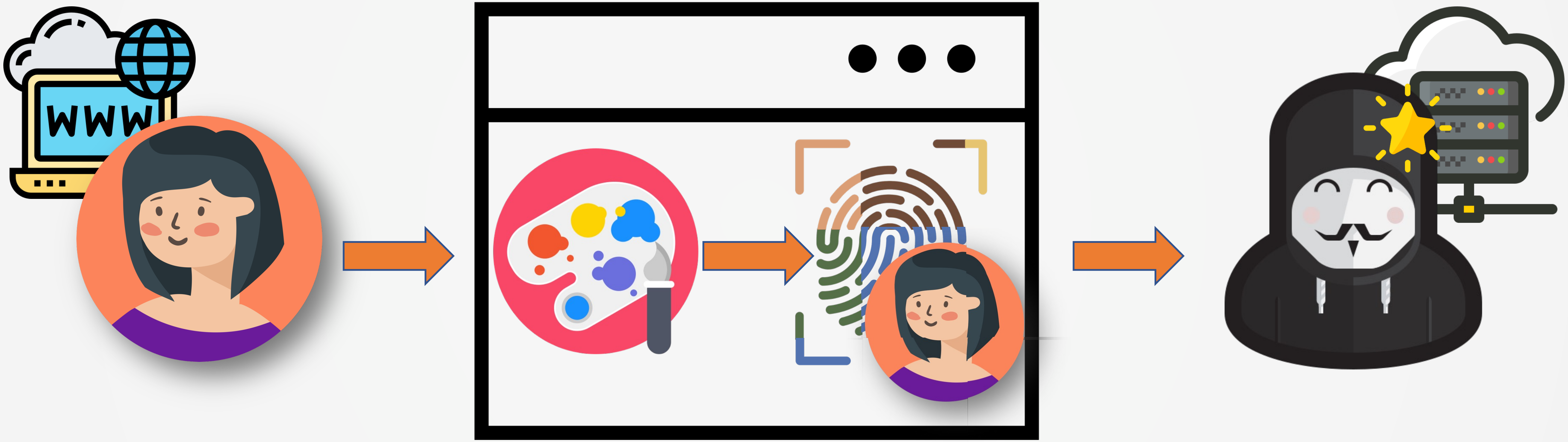
- ➢ Spoof certain APIs
- ➢ Disable JavaScript (entirely or partially)

Is fingerprinting possible without JavaScript?





Our approach: Implicit stylistic browser fingerprinting



Implicit stylistic browser fingerprinting

- ➢ Does not use any JavaScript
- ➢ Provides highly discriminating fingerprints

What can we use to detect the stylistic differences?

Chrome	Firefox	Safari
40px/**15**px	**183**px/**16**px	**34**px/**13**px

12-hour Time	24-hour Time
146px/32px	**99**px/32px

English OS	Chinese OS
425px/**35**px	425px/**41**px

42

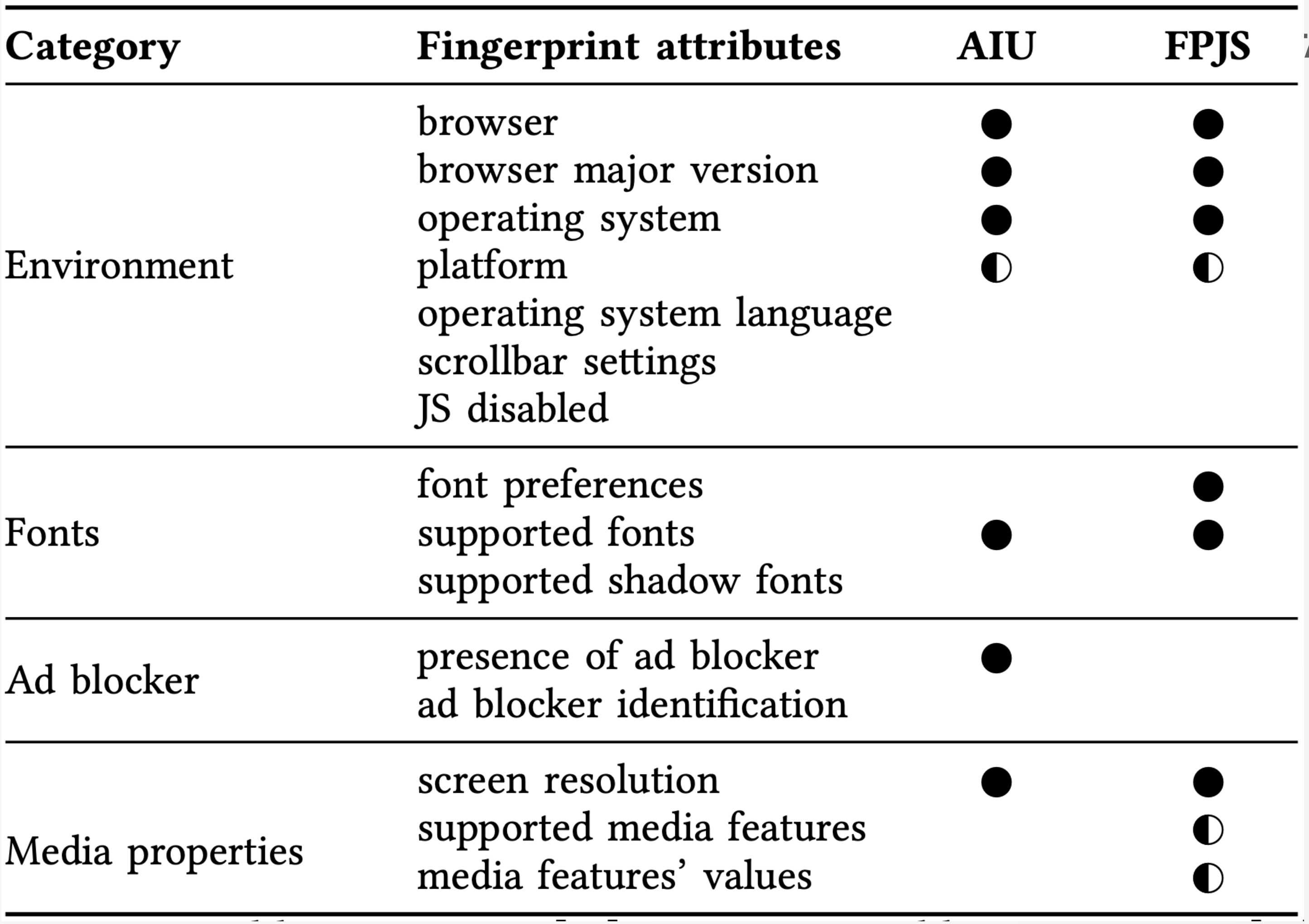
Dimensions!

Fingerprinting attributes

Certain HTML elements have different sizes depending on certain environmental factors.

339
Fingerprinting Elements

dimensions

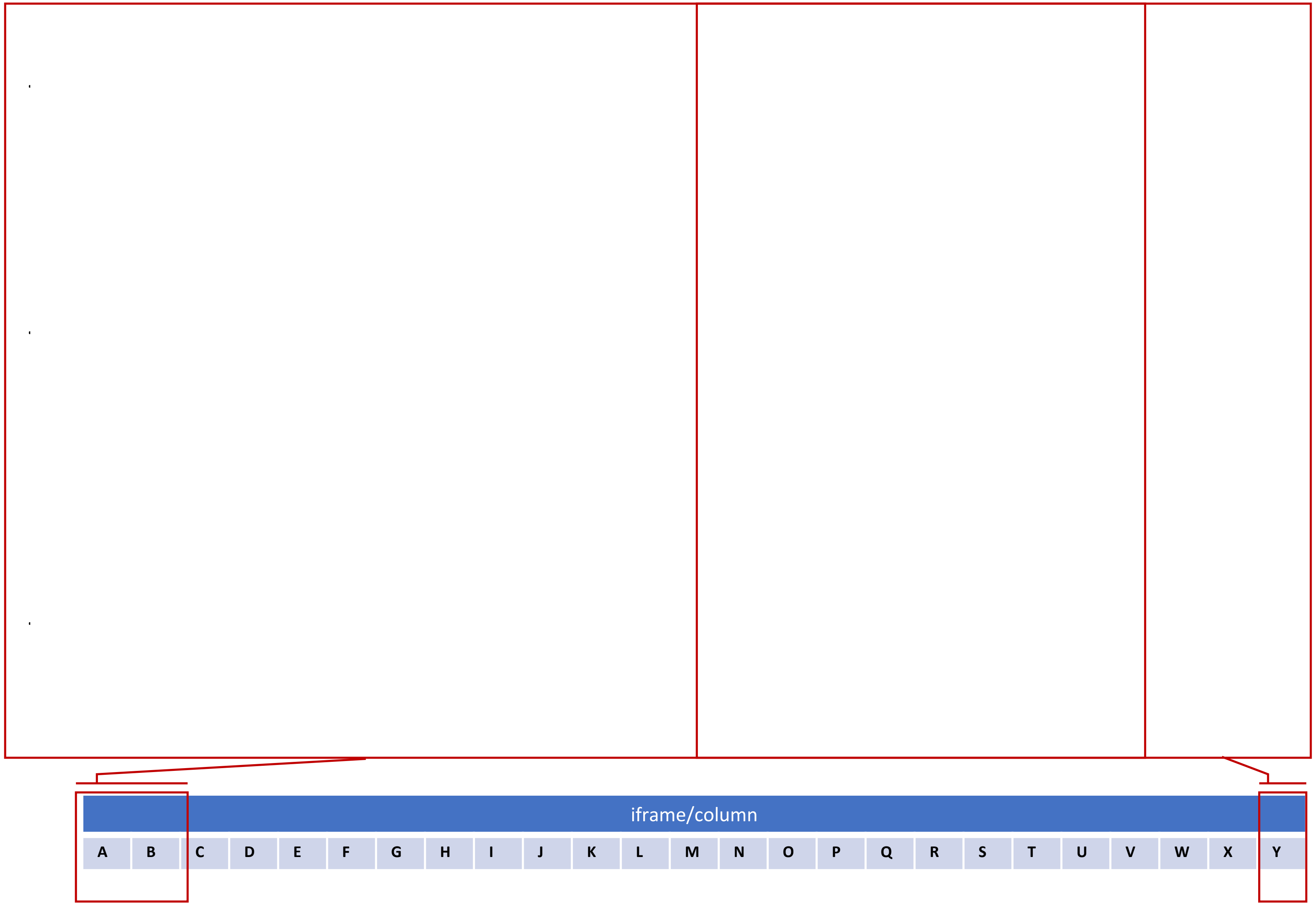


Category	Fingerprint attributes	AIU	FPJS
Environment	browser	●	●
	browser major version	●	●
	operating system	●	●
	platform	◐	◐
	operating system language		
	scrollbar settings		
	JS disabled		
Fonts	font preferences		●
	supported fonts	●	●
	supported shadow fonts		
Ad blocker	presence of ad blocker	●	
	ad blocker identification		
Media properties	screen resolution	●	●
	supported media features		◐
	media features' values		◐

AIU: captured by AmIUnique FPJS: captured by FingerprintJS

◐ : partial feature support ●: full feature support

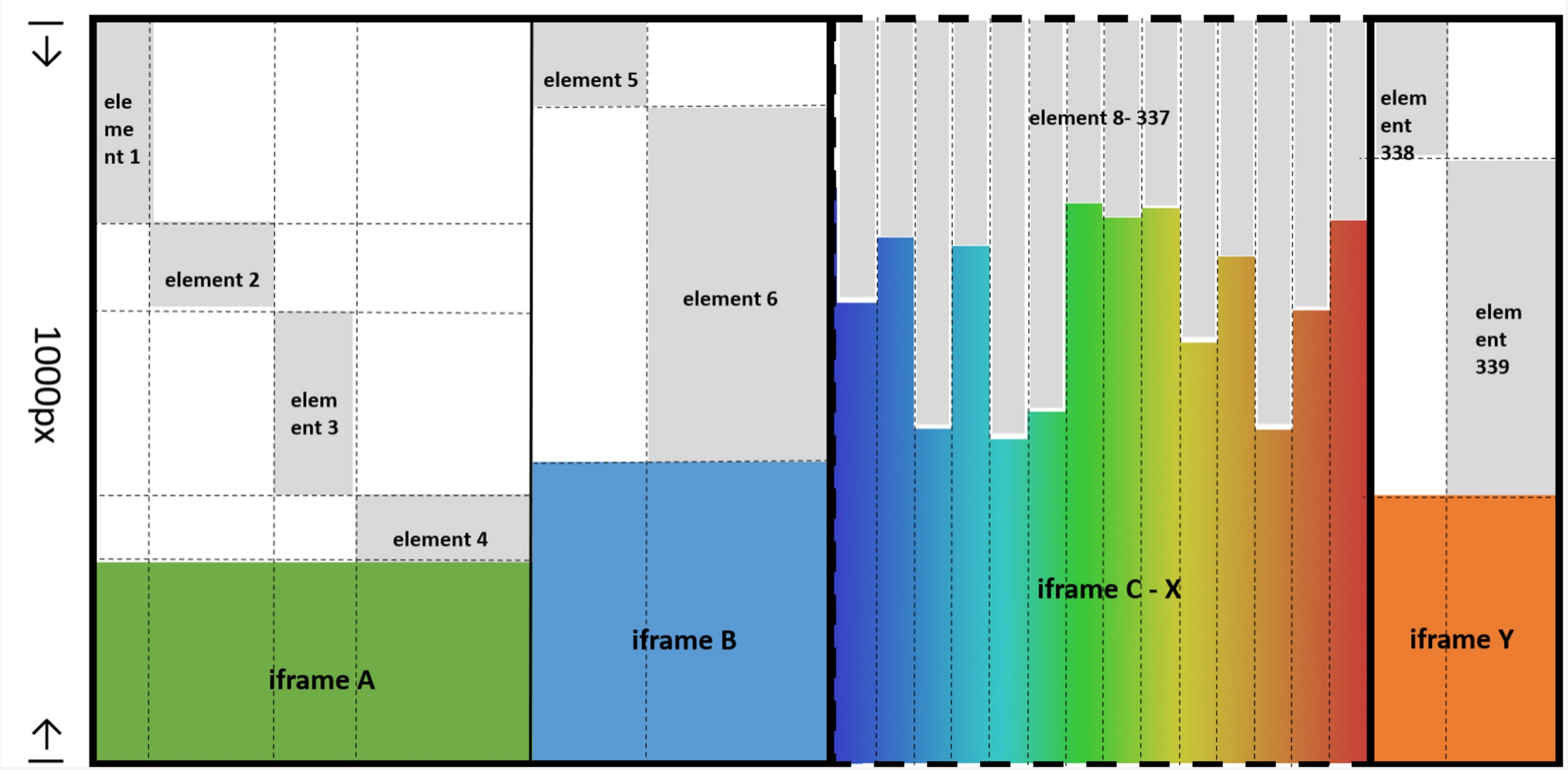
- The page only needs **25** iframes.
- All elements are placed in an 800px by 1000px iframe (main iframe) to ensure that their dimensions remain consistent across different browser window sizes.

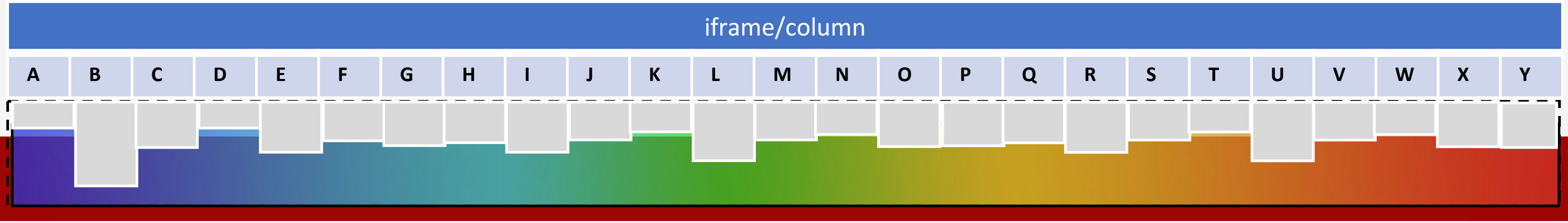


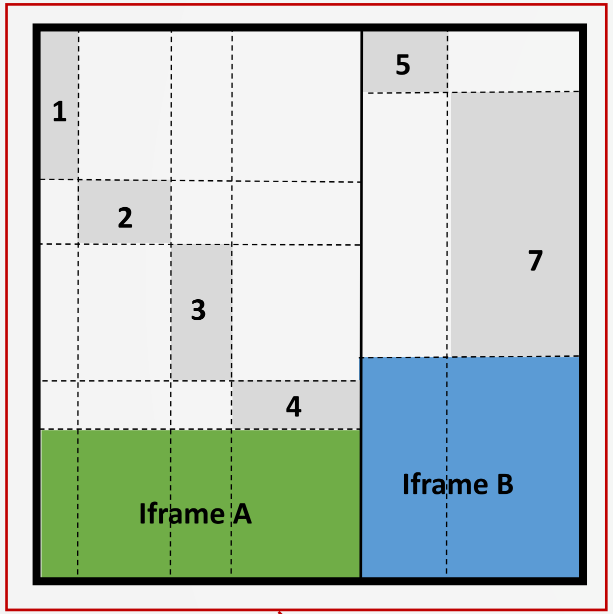
iframe/column

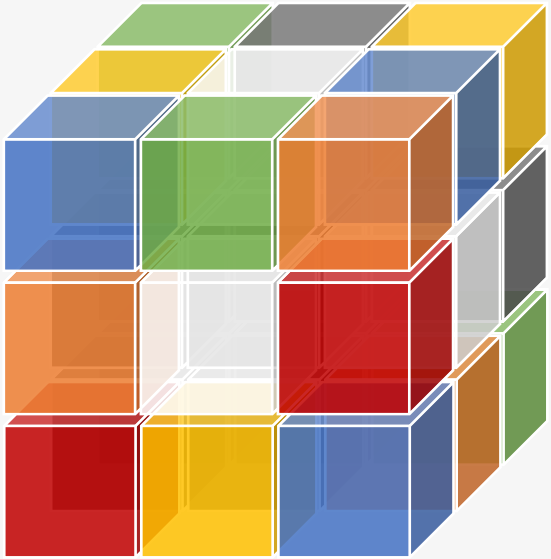
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

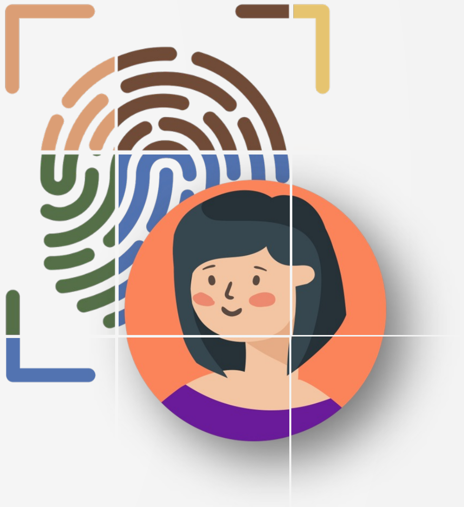
Main iframe

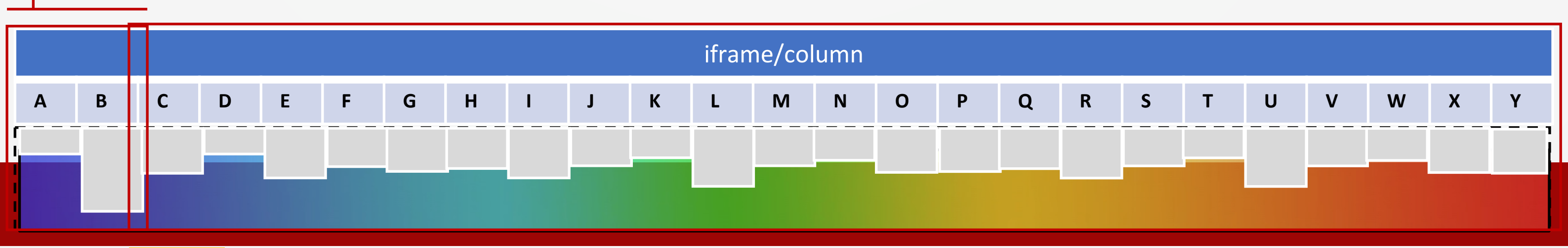

1000px
element 1
element 2
element 3
element 4
iframe A
element 5
element 6
iframe B
element 8- 337
iframe C - X
element 338
element 339
iframe Y


iframe/column
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y


1
2
3
4
5
7
Iframe A
Iframe B






iframe/column
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y

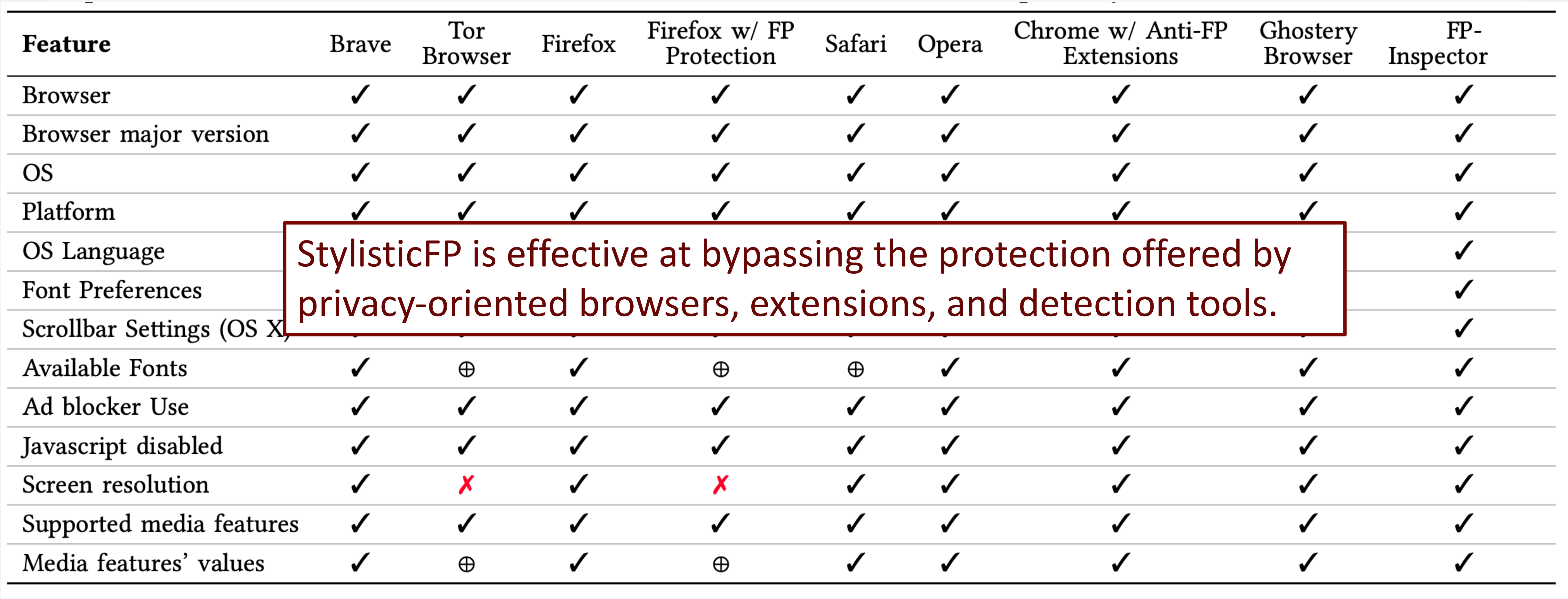
Evaluation

EFFECTIVENESS AGAINST ANTI-FINGERPRINTING BROWSERS AND TOOLS

CAPABILITY TO IDENTIFY DEVICES

Stylistic FP features effectiveness against popular countermeasures

✓denotes that our technique is effective, ✗ denotes that it is ineffective, and ⊕ denotes that it is partially effective.



Feature	Brave	Tor Browser	Firefox	Firefox w/ FP Protection	Safari	Opera	Chrome w/ Anti-FP Extensions	Ghostery Browser	FP-Inspector
Browser	✓	✓	✓	✓	✓	✓	✓	✓	✓
Browser major version	✓	✓	✓	✓	✓	✓	✓	✓	✓
OS	✓	✓	✓	✓	✓	✓	✓	✓	✓
Platform	✓	✓	✓	✓	✓	✓	✓	✓	✓
OS Language	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	✓
Font Preferences	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	✓
Scrollbar Settings (OS X)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	✓
Available Fonts	✓	⊕	✓	⊕	⊕	✓	✓	✓	✓
Ad blocker Use	✓	✓	✓	✓	✓	✓	✓	✓	✓
Javascript disabled	✓	✓	✓	✓	✓	✓	✓	✓	✓
Screen resolution	✓	✗	✓	✗	✓	✓	✓	✓	✓
Supported media features	✓	✓	✓	✓	✓	✓	✓	✓	✓
Media features' values	✓	⊕	✓	⊕	✓	✓	✓	✓	✓

StylisticFP is effective at bypassing the protection offered by privacy-oriented browsers, extensions, and detection tools.

We shared the source code and paper with browser venders upon requests, and received a bounty from Brave.

Evaluation

EFFECTIVENESS AGAINST ANTI-FINGERPRINTING BROWSERS AND TOOLS

CAPABILITY TO IDENTIFY DEVICES

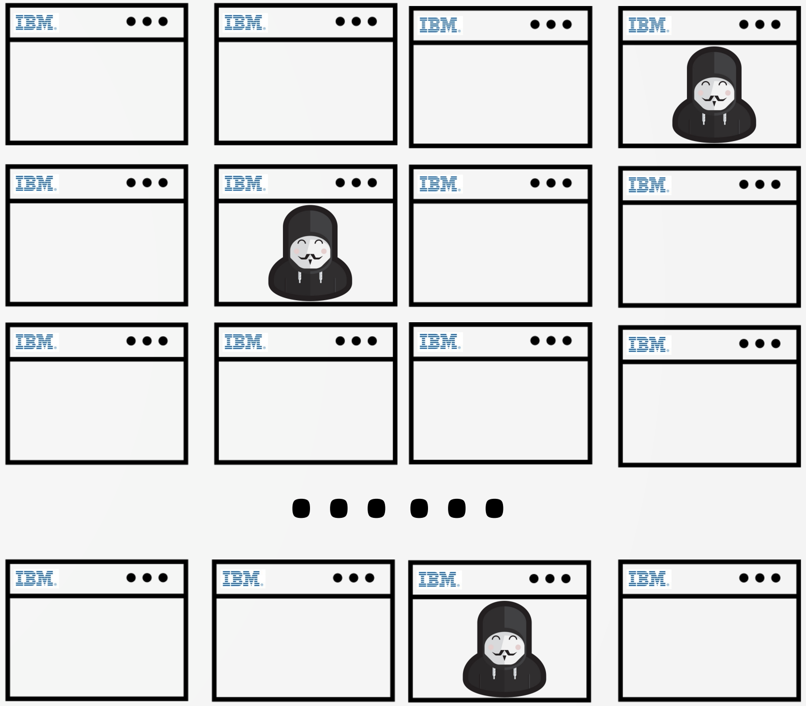
Pilot study

- Compared to FingerprintJS (FPJS), a prevalent state-of-the-art fingerprinting library.

- Systems deployed on three IBM intranet portals between June 1st – Aug 8th 2022”.

```
<iframe src="fp.url" style="visibility:hidden;"/>
```

- Device population is heavily skewed towards more specific, homogeneous models.



Capability to identify devices

StylisticFP

- possesses sufficient discriminative power
- outperforms FPJS in privacy-oriented browsers

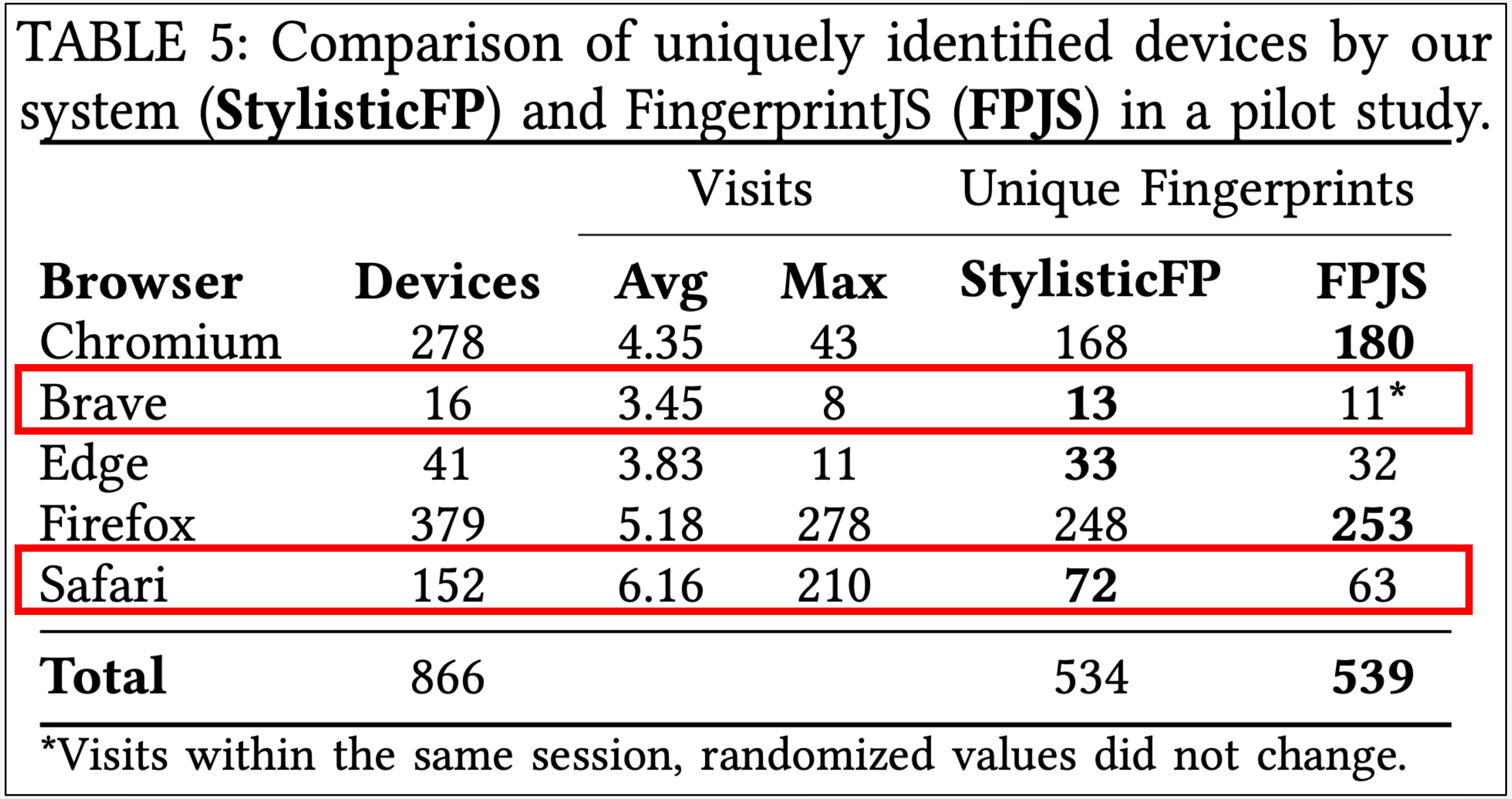


TABLE 5: Comparison of uniquely identified devices by our system (**StylisticFP**) and FingerprintJS (**FPJS**) in a pilot study.

		Visits		Unique Fingerprints	
Browser	**Devices**	**Avg**	**Max**	**StylisticFP**	**FPJS**
Chromium	278	4.35	43	168	**180**
Brave	16	3.45	8	**13**	11*
Edge	41	3.83	11	**33**	32
Firefox	379	5.18	278	248	**253**
Safari	152	6.16	210	**72**	63
Total	866			534	**539**

*Visits within the same session, randomized values did not change.

Our fingerprinting system

No JavaScript needed.

Comparable discriminating power to FPJS.

Effectively bypasses state-of-the-art anti-fingerprinting defenses.

If you're interested in breaking and fixing on the Web, please contact me at: xu.lin@wsu.edu