Building User-Centred Privacy Enhancing Technologies

Hamed Haddadi

PETS 2024, Bristol, UK

IMPERIAL

brave

Can we build **trusted**, *scalable,* **human-centred** systems:

… to perform *accurate* and *personalized* analytics;

… across the variety of ambient and personal data;

… **without** jeopardising the individuals’ *privacy*, *security*?

FACT

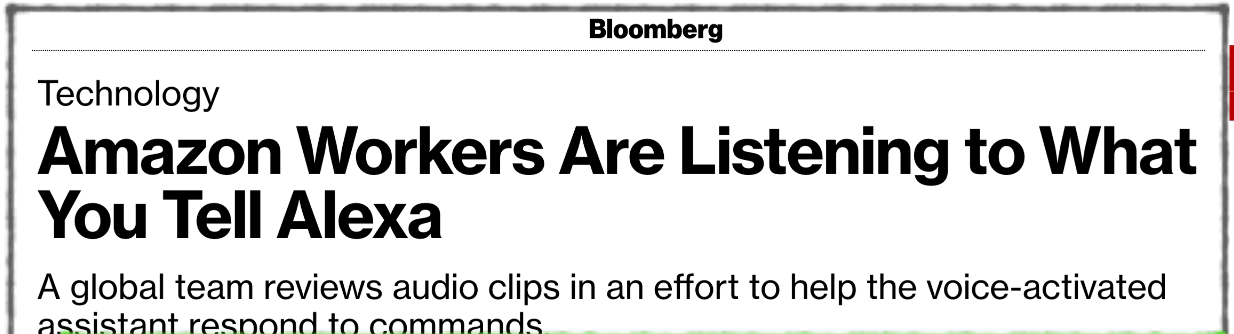
Part 1: IoT data

Data-Driven Networked Systems

They may listen to you
(e.g., smart speakers)

They may watch you
(e.g., smart doorbells)

They may know what
you watch (e.g., smart TVs)

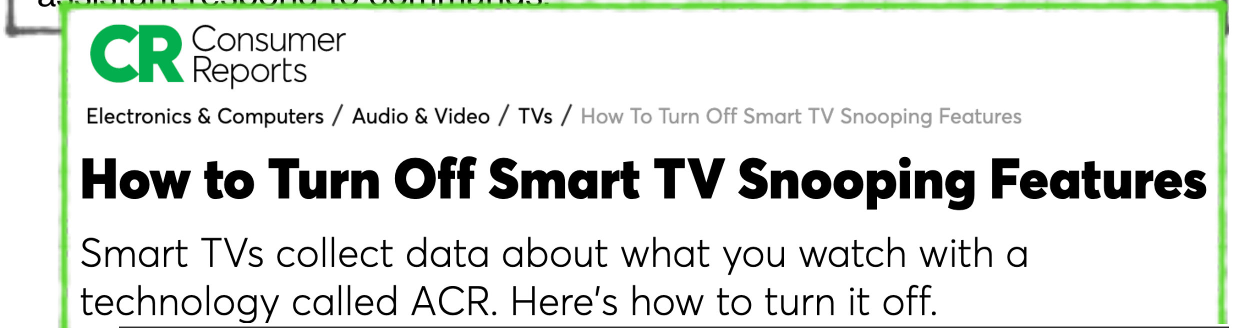


Bloomberg

Technology

Amazon Workers Are Listening to What You Tell Alexa

A global team reviews audio clips in an effort to help the voice-activated assistant respond to commands.

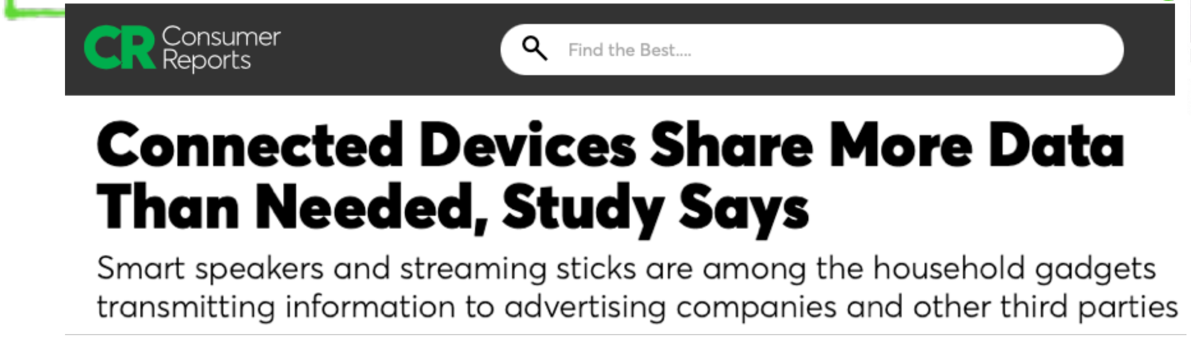


CR Consumer Reports

Electronics & Computers / Audio & Video / TVs / How To Turn Off Smart TV Snooping Features

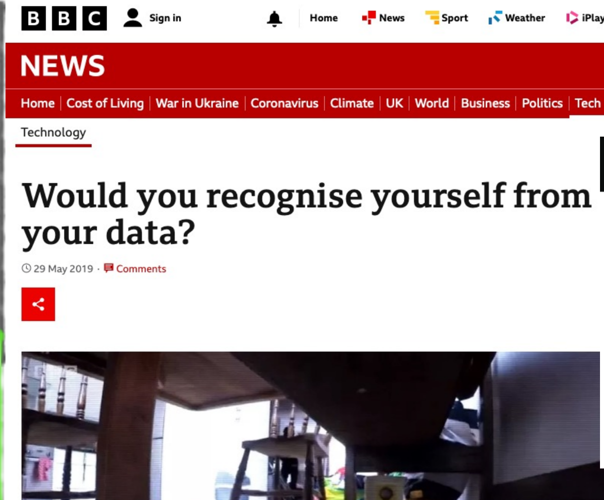
How to Turn Off Smart TV Snooping Features

Smart TVs collect data about what you watch with a technology called ACR. Here's how to turn it off.



Connected Devices Share More Data Than Needed, Study Says

Smart speakers and streaming sticks are among the household gadgets transmitting information to advertising companies and other third parties



BBC
Sign in
Home
News
Sport
Weather
iPlayer
NEWS
Home | Cost of Living | War in Ukraine | Coronavirus | Climate | UK | World | Business | Politics | Tech

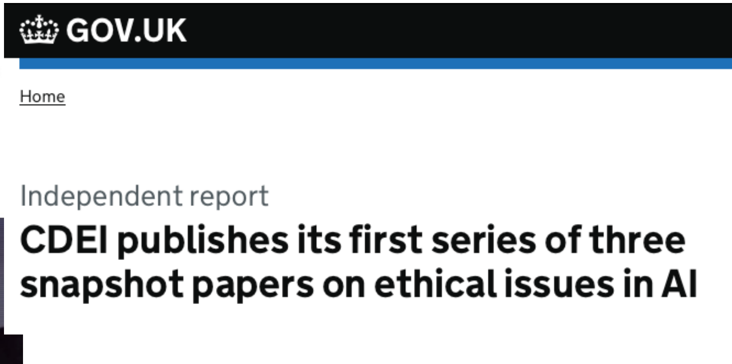
Technology

Would you recognise yourself from your data?

29 May 2019 · Comments

Carl's vacuum's view of his house

By Carl Miller
Research director and author, Demos



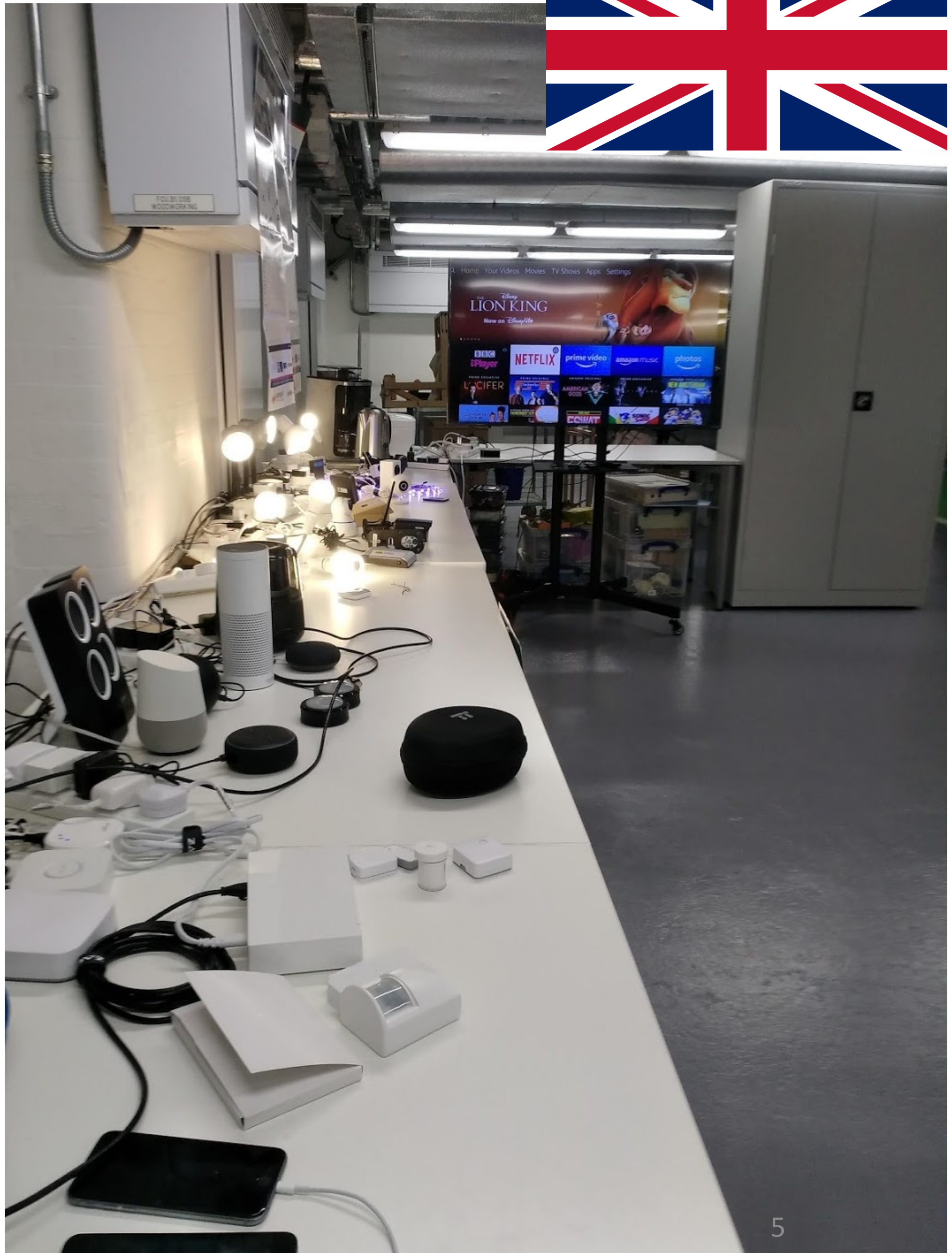
Independent report

CDEI publishes its first series of three snapshot papers on ethical issues in AI

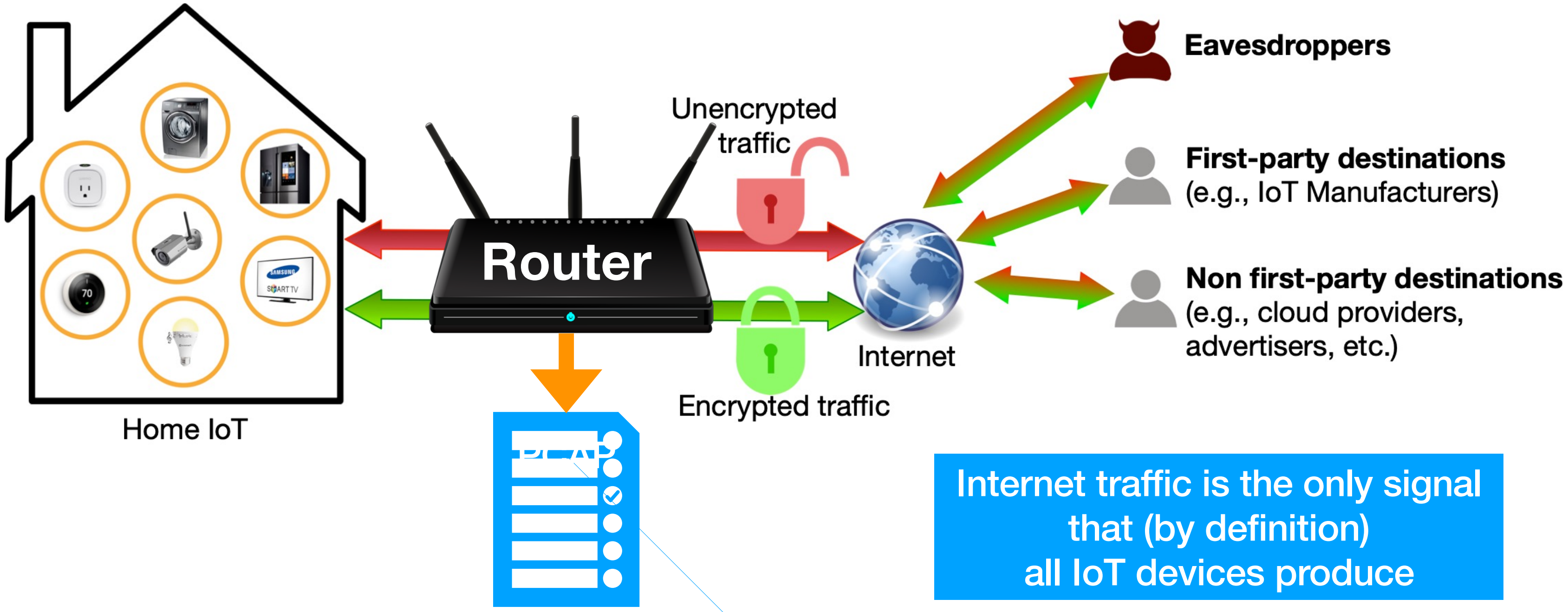
The Truth About Amazon

As the high street goes into lockdown, Amazon is booming. This Supershoppers special reveals how to buy smart off the online retail giant, from the best bargains to avoiding scams.

140+ devices in two different countries



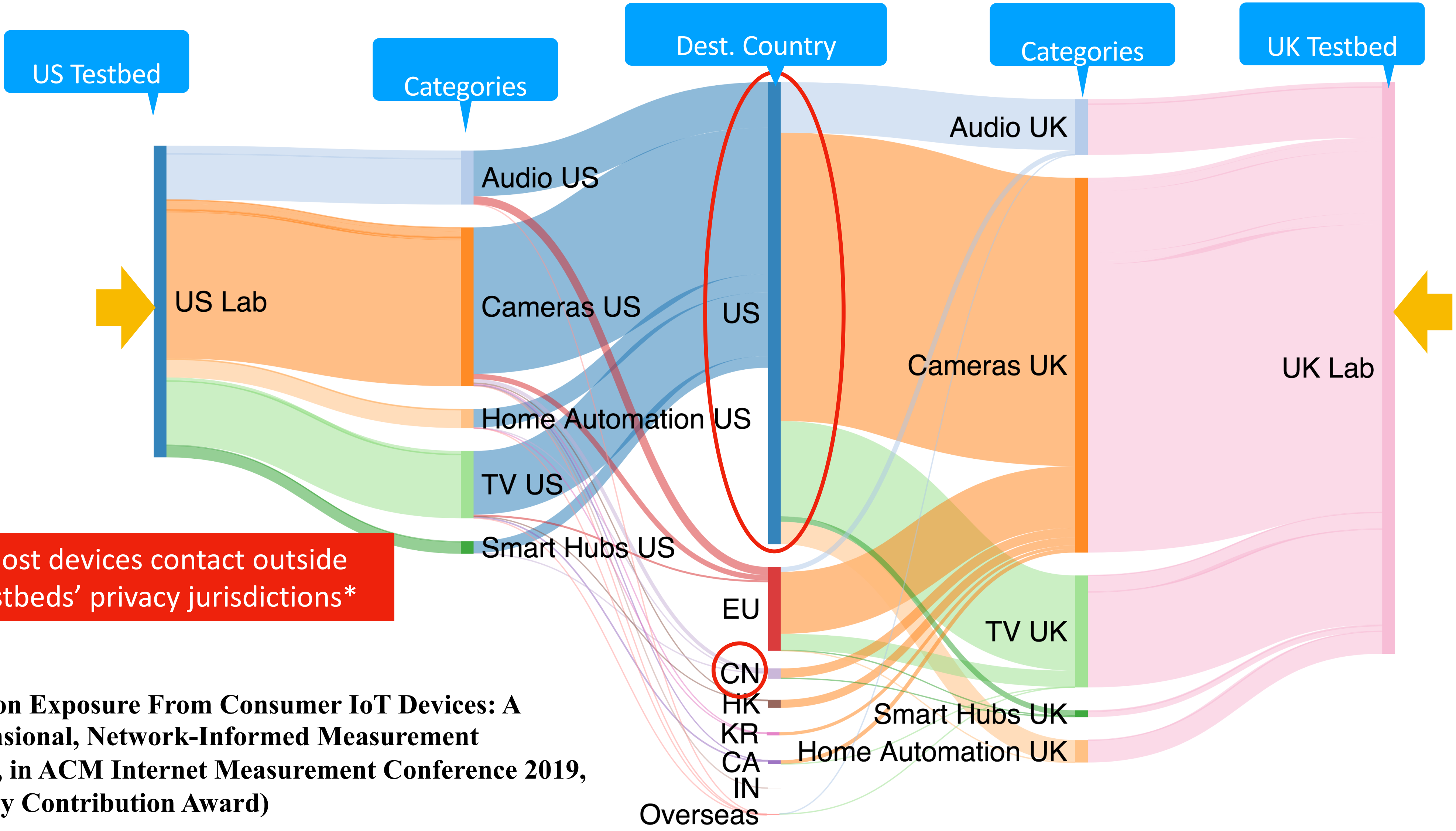
Data Collection Methodology


Home IoT
SAMSUNG
SMART TV
70
Router
PCAP
Unencrypted traffic
Encrypted traffic
Internet
Eavesdroppers
First-party destinations
(e.g., IoT Manufacturers)
Non first-party destinations
(e.g., cloud providers, advertisers, etc.)

Internet traffic is the only signal that (by definition) all IoT devices produce

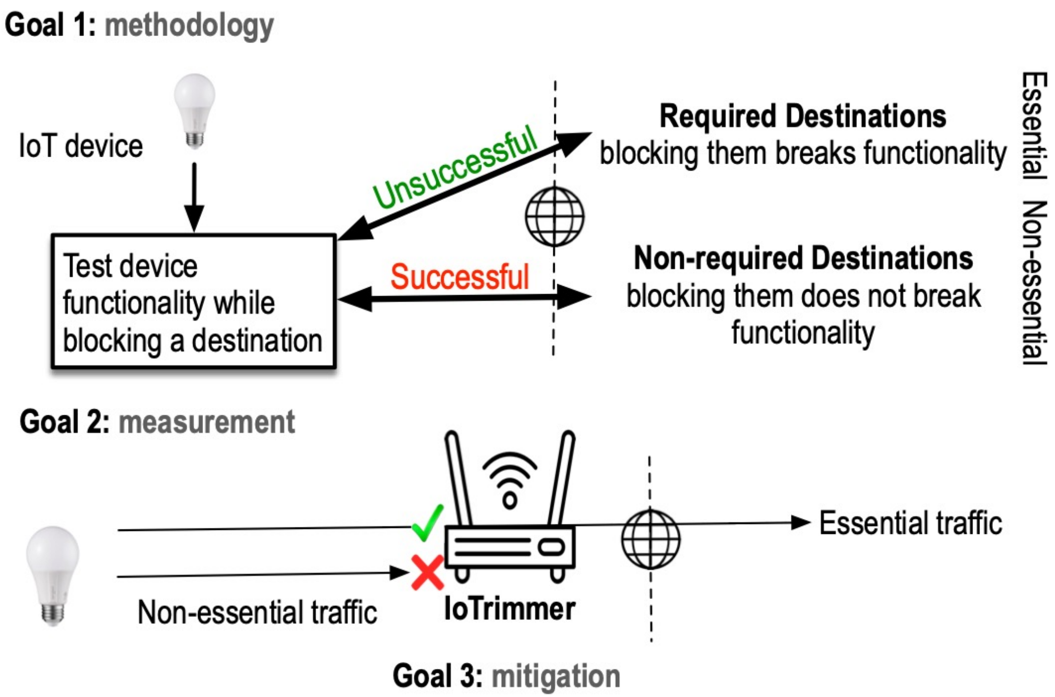
- Monitor all traffic at the **router**
  - per-device
  - per-experiment

Most traffic goes beyond Europe



Most devices contact outside testbeds' privacy jurisdictions*

"Information Exposure From Consumer IoT Devices: A Multidimensional, Network-Informed Measurement Approach", in ACM Internet Measurement Conference 2019, (Community Contribution Award)



Blocking without Breaking

PETS 2021, Oakland 2023

iotrim.net

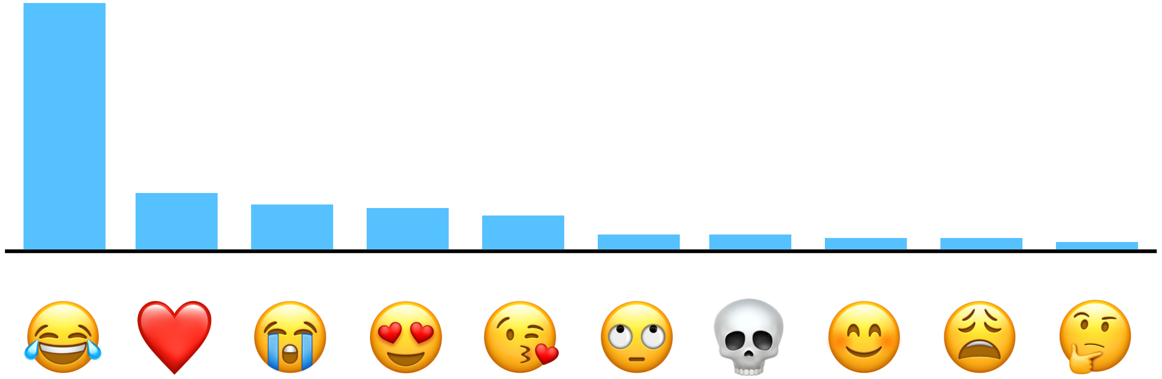
We squeezed more ML into routers..

Then came the regulators, governments, and the cops…

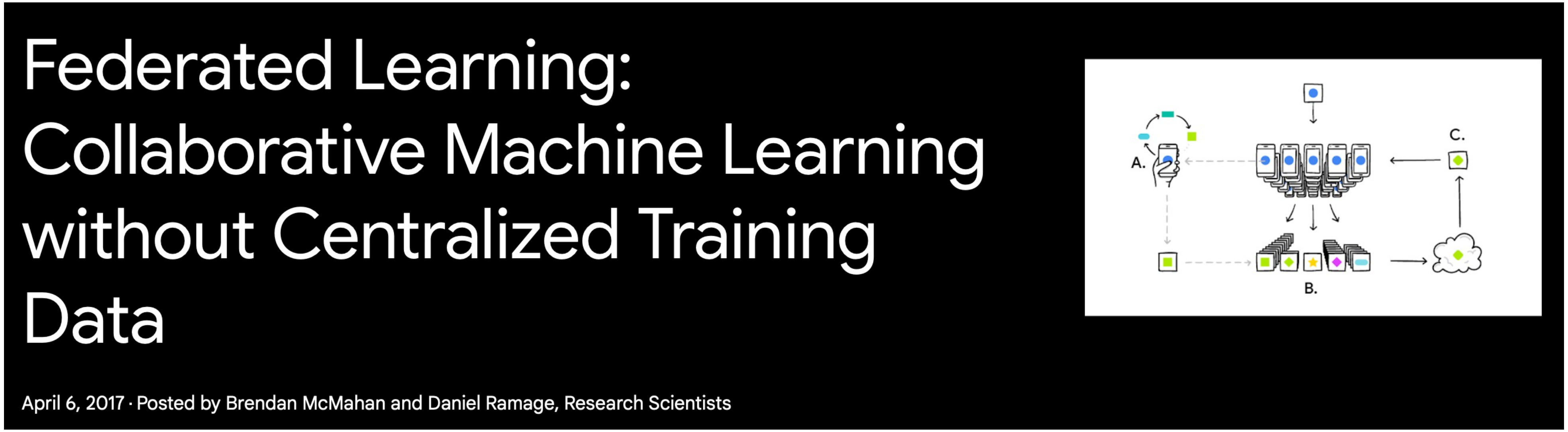
Everyone loves the smell of user data…

Part 2: Mobile & web data

Telemetry is becoming popular



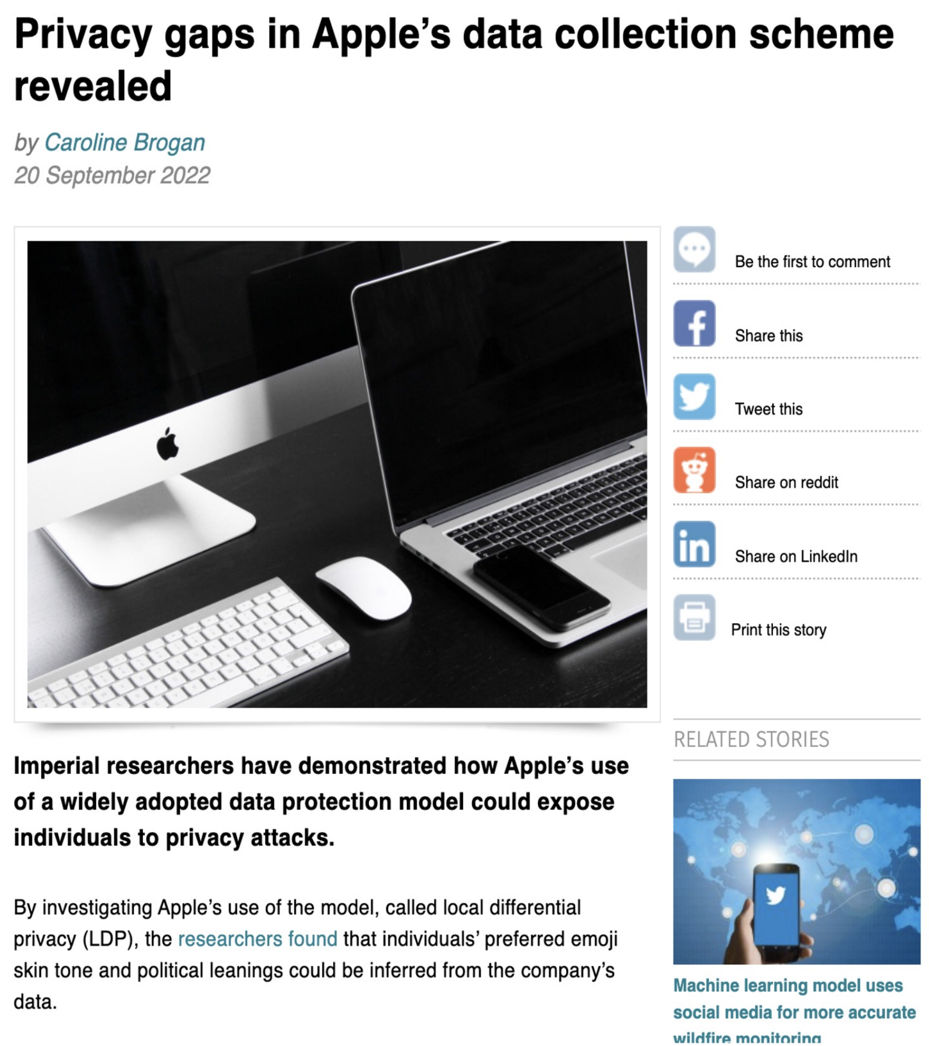
The Count Mean Sketch technique allows Apple to determine the most popular emoji to help design better ways to find and use our favorite emoji. The top emoji for US English speakers contained some surprising favorites.



Federated Learning: Collaborative Machine Learning without Centralized Training Data

April 6, 2017 · Posted by Brendan McMahan and Daniel Ramage, Research Scientists

But it comes at a cost



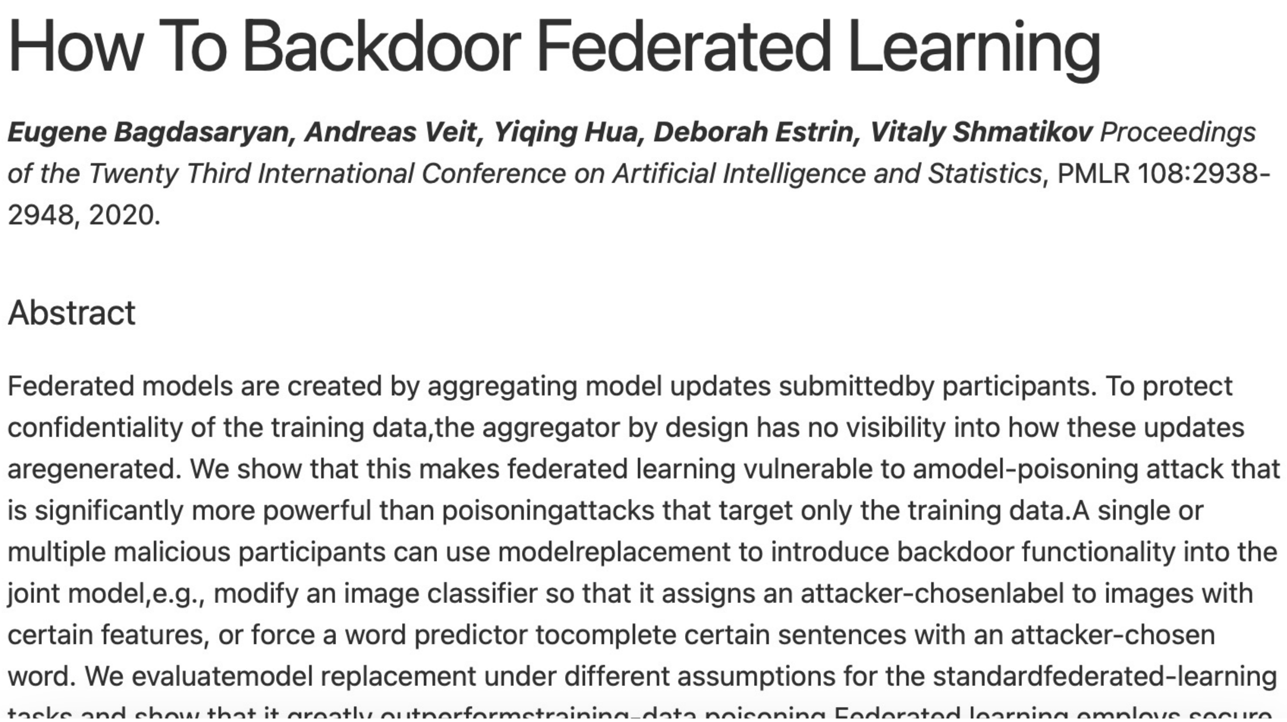
Privacy gaps in Apple's data collection scheme revealed

by Caroline Brogan
20 September 2022

Imperial researchers have demonstrated how Apple's use of a widely adopted data protection model could expose individuals to privacy attacks.

By investigating Apple's use of the model, called local differential privacy (LDP), the researchers found that individuals' preferred emoji skin tone and political leanings could be inferred from the company's data.

Machine learning model uses social media for more accurate wildfire monitoring



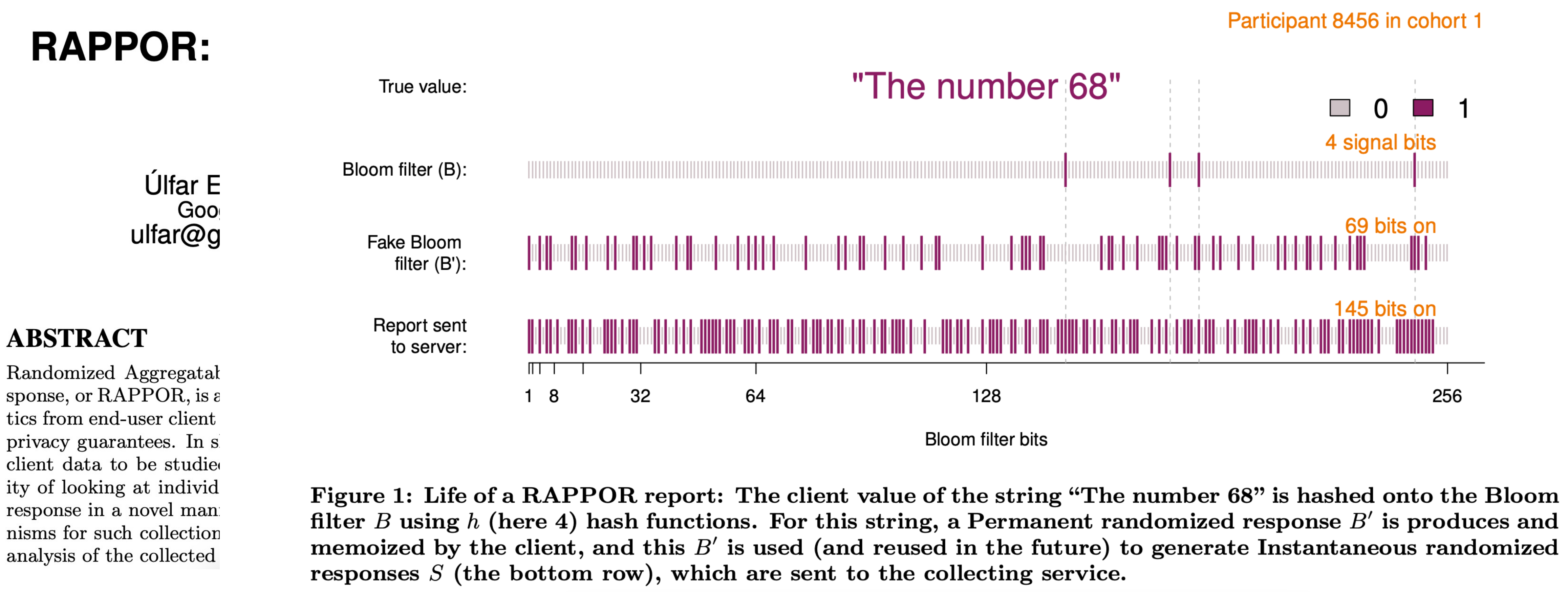
How To Backdoor Federated Learning

Eugene Bagdasaryan, Andreas Veit, Yiqing Hua, Deborah Estrin, Vitaly Shmatikov *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, PMLR 108:2938-2948, 2020.

Abstract

Federated models are created by aggregating model updates submittedby participants. To protect confidentiality of the training data,the aggregator by design has no visibility into how these updates aregenerated. We show that this makes federated learning vulnerable to amodel-poisoning attack that is significantly more powerful than poisoningattacks that target only the training data.A single or multiple malicious participants can use modelreplacement to introduce backdoor functionality into the joint model,e.g., modify an image classifier so that it assigns an attacker-chosenlabel to images with certain features, or force a word predictor tocomplete certain sentences with an attacker-chosen word. We evaluatemodel replacement under different assumptions for the standardfederated-learning tasks and show that it greatly outperformstraining-data poisoning.Federated learning employs secure

Browser telemetry, Rappor



RAPPOR:

Úlfar E
Goo
ulfar@g

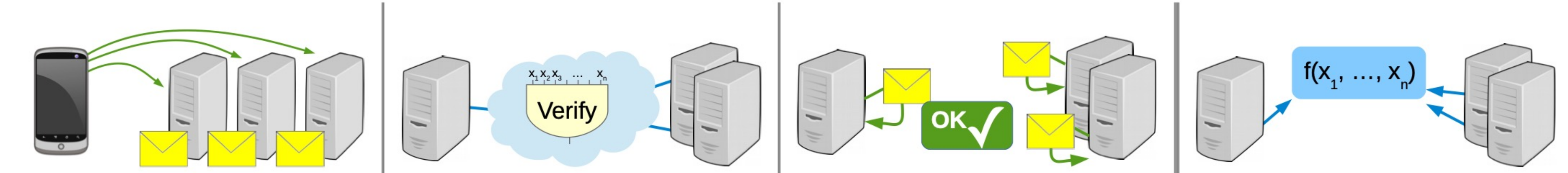
ABSTRACT

Randomized Aggregatal
sponse, or RAPPOR, is a
tics from end-user client
privacy guarantees. In s
client data to be studie
ity of looking at individ
response in a novel man
nisms for such collection
analysis of the collected

Figure 1: Life of a RAPPOR report: The client value of the string "The number 68" is hashed onto the Bloom filter B using h (here 4) hash functions. For this string, a Permanent randomized response B' is produces and memoized by the client, and this B' is used (and reused in the future) to generate Instantaneous randomized responses S (the bottom row), which are sent to the collecting service.

Browser telemetry, Prio

Mozilla Security Blog

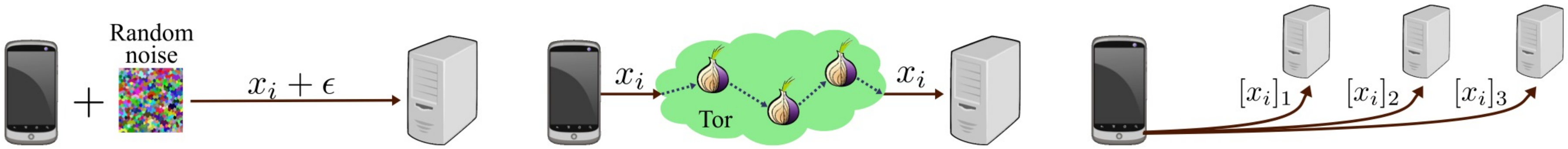


(a) The client sends a share of its encoded submission and SNIP proof to each server.

(b) The servers validate the client's SNIP proof to ensure that the submission is valid.

(c) If the checks pass, the servers update their local accumulators with the client-provided data.

(d) After accumulating many packets, the servers publish their accumulators to reveal the aggregate.



(a) RAPPOR [57] provides differential privacy [54] (not information-theoretic privacy) by adding random noise to client submissions.

(b) ANONIZE [76] and PrivStats [100] rely on an anonymizing proxy, such as Tor [51], to protect privacy against network eavesdroppers.

(c) Prio and other schemes using secret sharing [30,48,56,79,86,92] offer ideal anonymity provided that the servers do not collude.

Figure 10: Comparison of techniques for anonymizing client data in private aggregation systems.

Browser telemetry, Dprio, Prio+

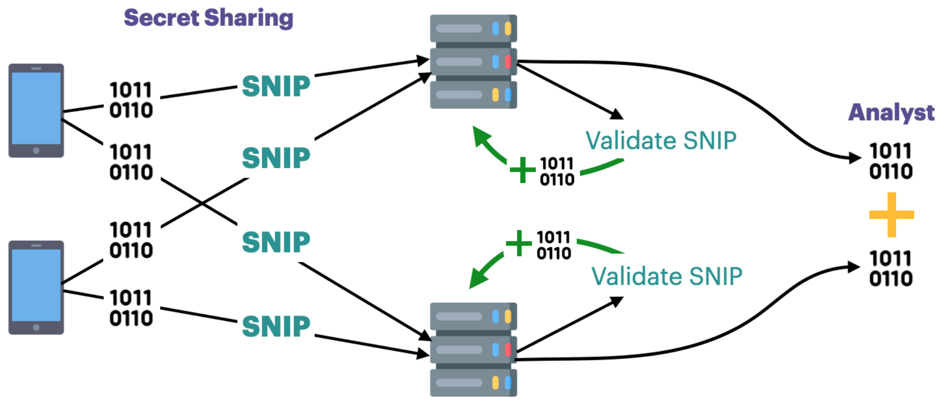
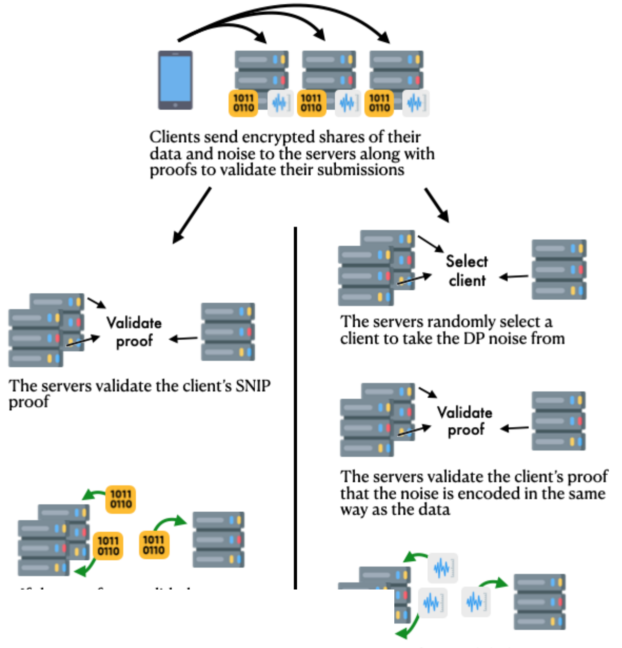


Figure 1: Overview of Prio. Clients send shares to servers who validate the associated SNIP and aggregate the data.



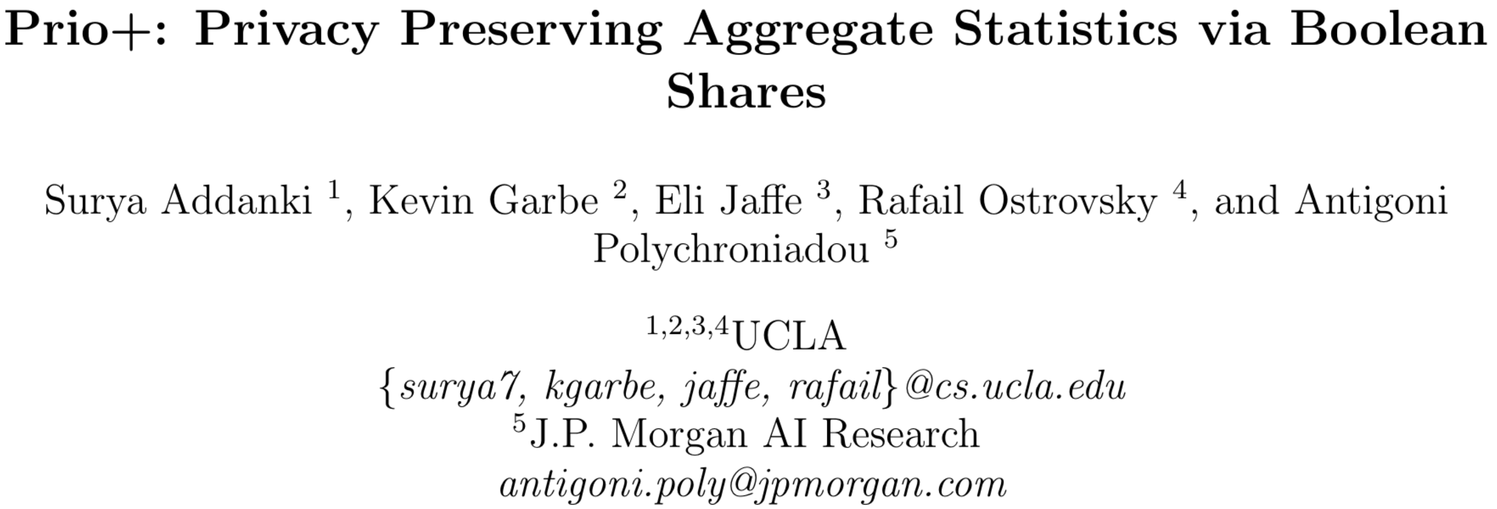
Clients send encrypted shares of their data and noise to the servers along with proofs to validate their submissions

The servers validate the client's SNIP proof

The servers randomly select a client to take the DP noise from

The servers validate the client's proof that the noise is encoded in the same way as the data

roofs are valid, the
update their aggregators
noise



Prio+: Privacy Preserving Aggregate Statistics via Boolean Shares

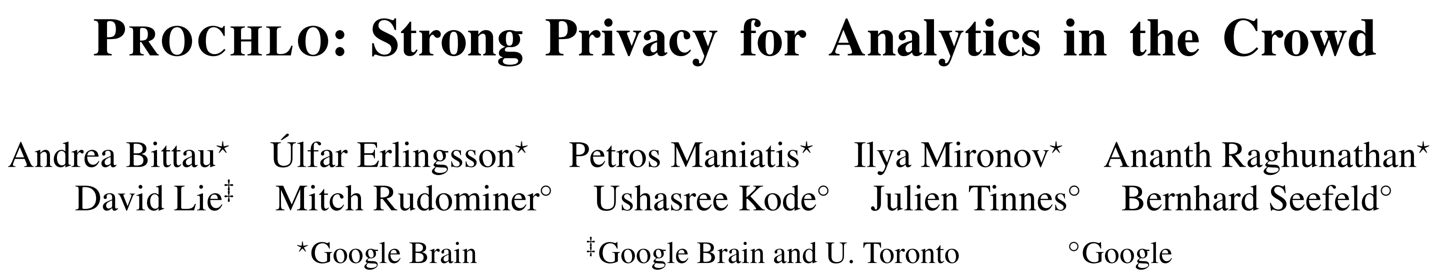
Surya Addanki [1], Kevin Garbe [2], Eli Jaffe [3], Rafail Ostrovsky [4], and Antigoni Polychroniadou [5]

[1,2,3,4]UCLA
{surya7, kgarbe, jaffe, rafail}@cs.ucla.edu
[5]J.P. Morgan AI Research
antigoni.poly@jpmorgan.com

are of
ults
lt

**s send shares to servers
lect a client's noise, and
the noise.**

Browser telemetry, Prochlo



PROCHLO: Strong Privacy for Analytics in the Crowd

Andrea Bittau[⋆] Úlfar Erlingsson[⋆] Petros Maniatis[⋆] Ilya Mironov[⋆] Ananth Raghunathan[⋆]
David Lie[‡] Mitch Rudominer[○] Ushasree Kode[○] Julien Tinnes[○] Bernhard Seefeld[○]

[⋆]Google Brain [‡]Google Brain and U. Toronto [○]Google

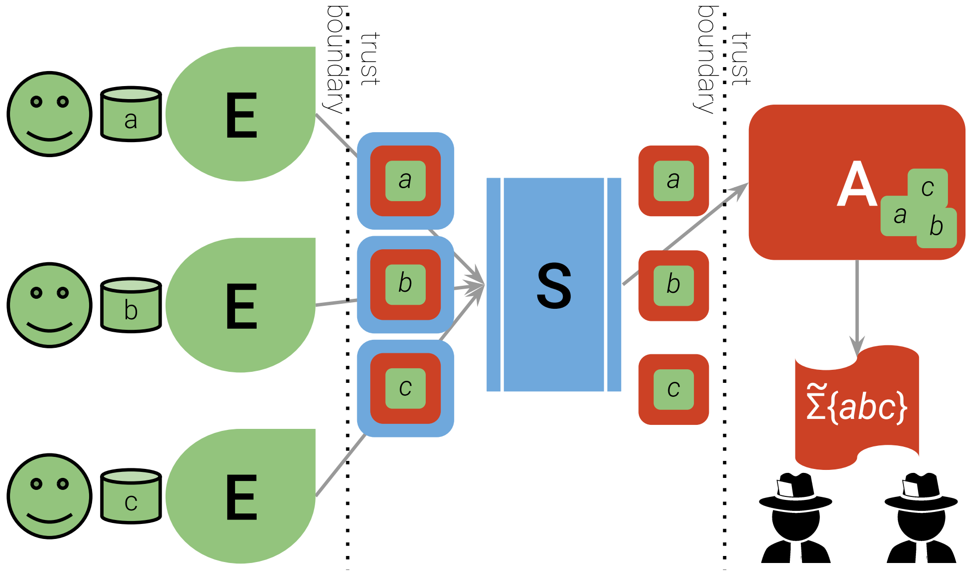
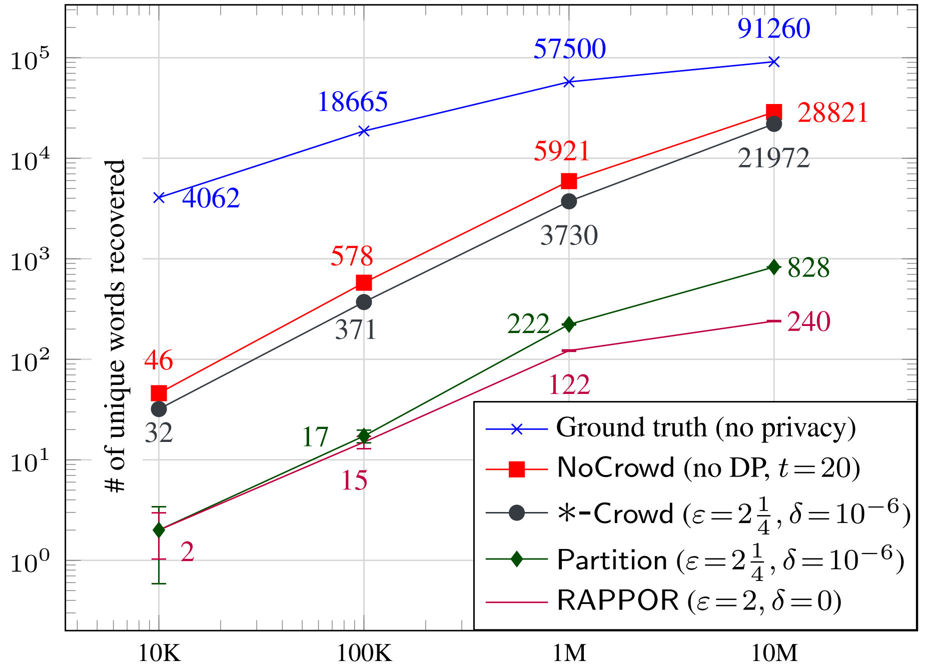


Figure 1: ESA architecture: Encode, shuffle, and analyze.



Browser telemetry, P3A

Privacy-Preserving Product Analytics (P3A)

PRIVACY UPDATES

STAR: Brave’s new system for privacy-preserving data collection

By the Brave Privacy Team

- “STAR: Secret Sharing for Private Threshold Aggregation Reporting”, ACM CCS 2022, ***Distinguished Paper Award***

Browser telemetry, STAR >> Nebula

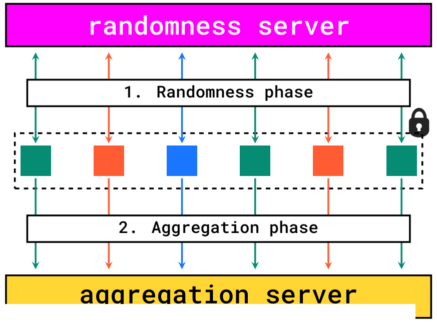
STAR: SECRET SHARING FOR THRESHOLD AGGREGATION REPORTING

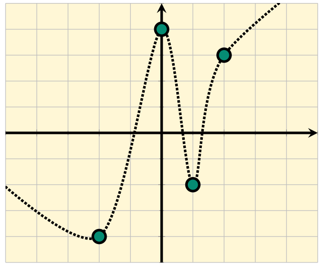
Alex Davidson[1] **Peter Snyder**[1] Joseph Genereux[1]
E. B. Quirk[1] Benjamin Livshits[2] Hamed Haddadi[1,2]

[1]Brave Software

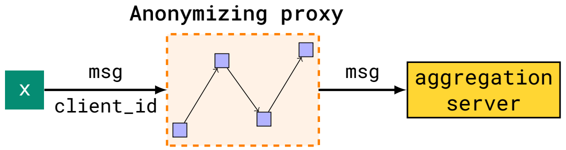
[2]Imperial College London

ACM CCS 2022 ::

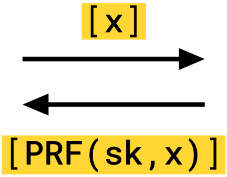




Shamir secret sharing



Anonymizing proxy (such as Tor, or Oblivious HTTP)



Oblivious PRF

c = Enc(ek,m)

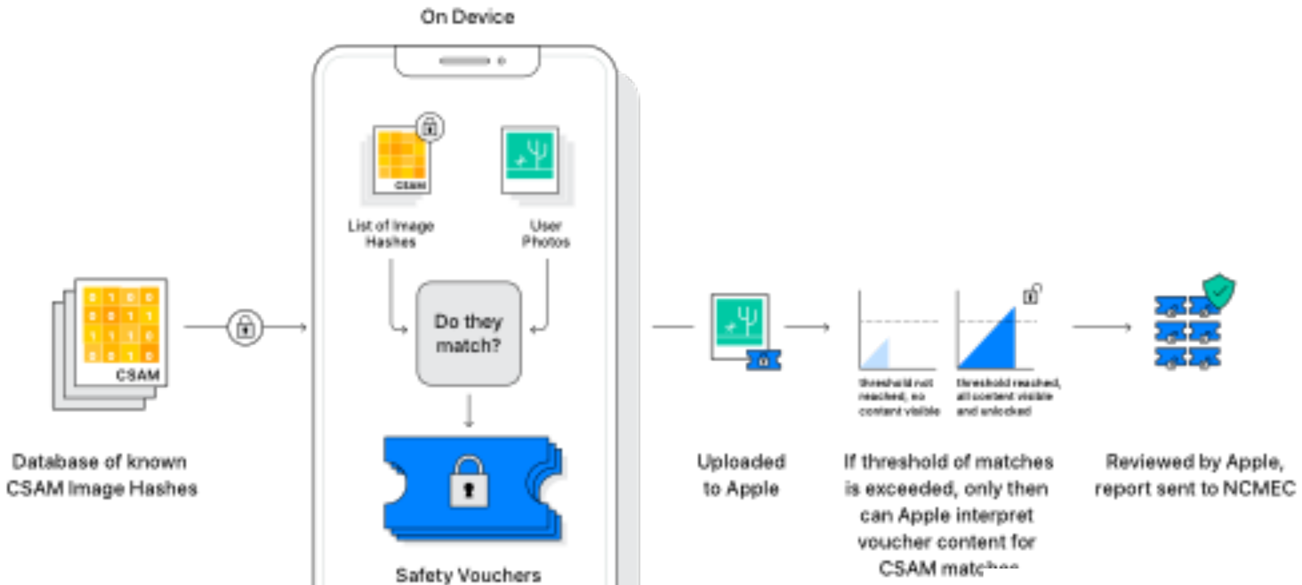
Symmetric encryption

- ◇ Emphasis on simplicity and performance
- ◇ Well-known cryptography (secret sharing, OPRFs)
- ◇ Orders of magnitude cheaper than state-of-the-art
- ◇ Malicious security
- ◇ Auxiliary data support
- ◇ Open-source rust code: github.com/brave/sta-rs

But, why do we need so many telemetry mechanisms?

- Different vendors, different requirements
- ***Greed over time***
- Will interoperability (think EU DMA) or standardisation efforts help?
- Will lobbying by the bigger forces prevent true privacy?

Other device analytics: CSAM



On Device

CSAM

List of Image Hashes

User Photos

Do they match?

Safety Vouchers

Database of known CSAM Image Hashes

Uploaded to Apple

threshold not reached, no content visible

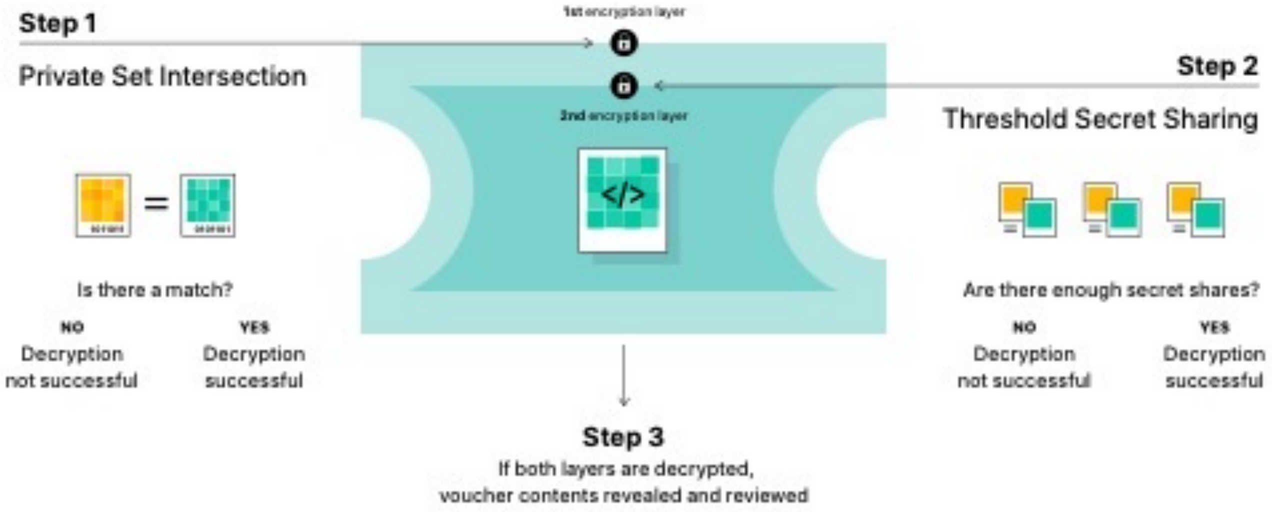
threshold reached, all content visible and unlocked

If threshold of matches is exceeded, only then can Apple interpret voucher content for CSAM matc

Reviewed by Apple, report sent to NCMEC

Expanded Protections for Children

August 2021



Step 1

Private Set Intersection

Is there a match?

NO
Decryption not successful

YES
Decryption successful

1st encryption layer

2nd encryption layer

Step 2

Threshold Secret Sharing

Are there enough secret shares?

NO
Decryption not successful

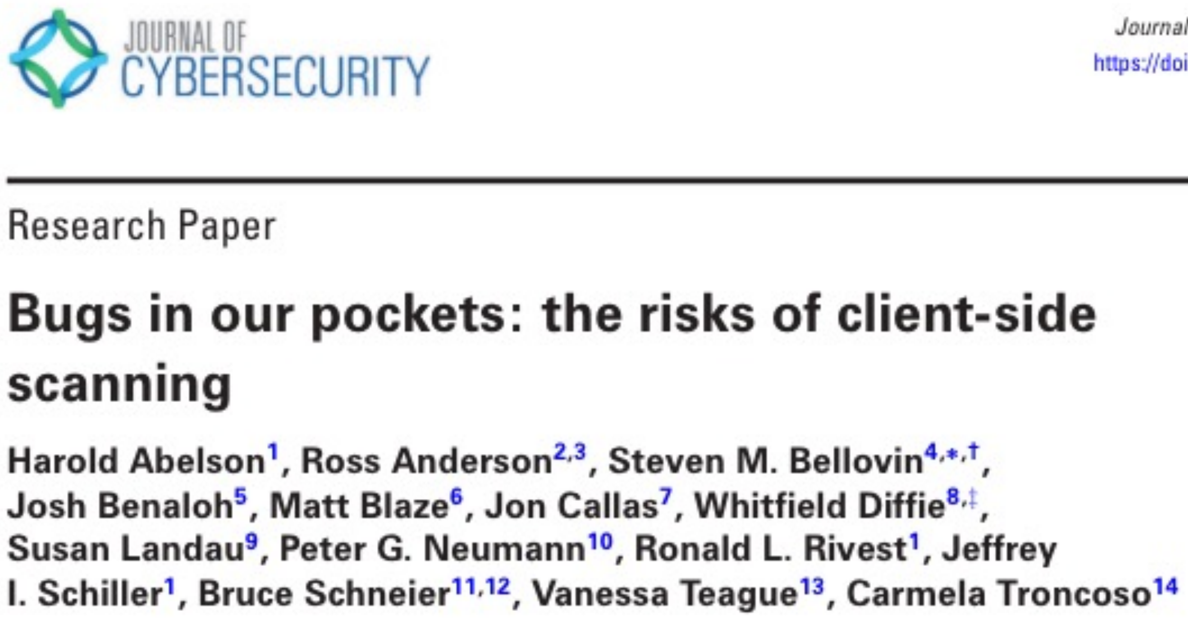
YES
Decryption successful

Step 3

If both layers are decrypted, voucher contents revealed and reviewed

system are described on the main project site.

CSAM failures



JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2024, 1–18
https://doi.org/10.1093/cybsec/tyad020

Research Paper

Bugs in our pockets: the risks of client-side scanning

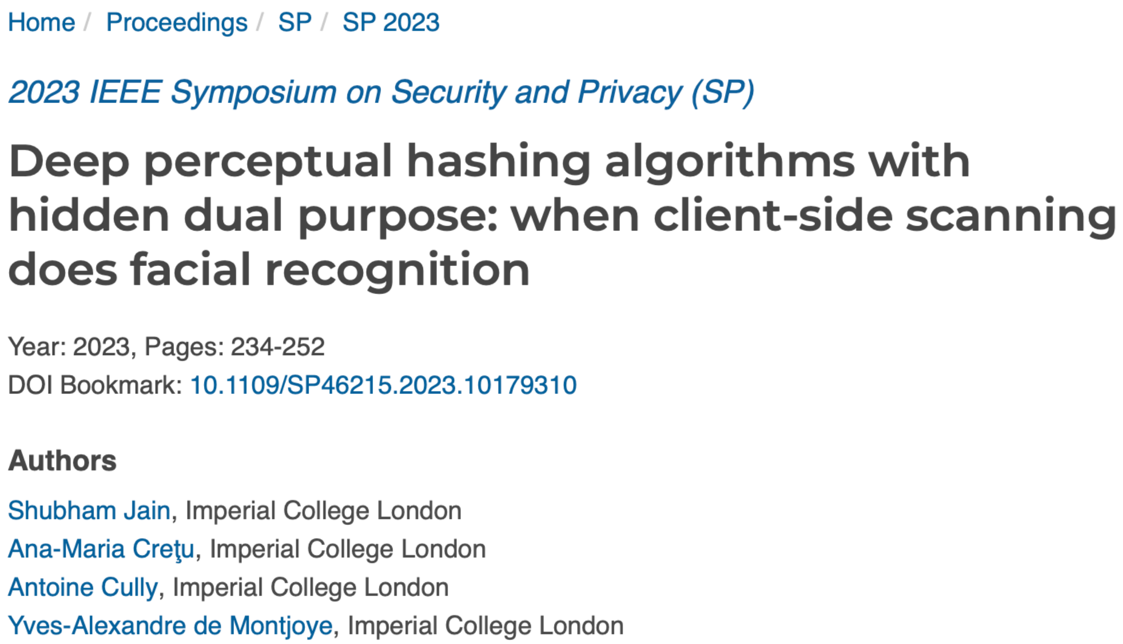
Harold Abelson[1], Ross Anderson[2,3], Steven M. Bellovin[4,*,†], Josh Benaloh[5], Matt Blaze[6], Jon Callas[7], Whitfield Diffie[8,‡], Susan Landau[9], Peter G. Neumann[10], Ronald L. Rivest[1], Jeffrey I. Schiller[1], Bruce Schneier[11,12], Vanessa Teague[13], Carmela Troncoso[14]

False positives

Collision attacks

Misuse by authoritarian governments

Potential expansion into messaging



Home / Proceedings / SP / SP 2023

2023 IEEE Symposium on Security and Privacy (SP)

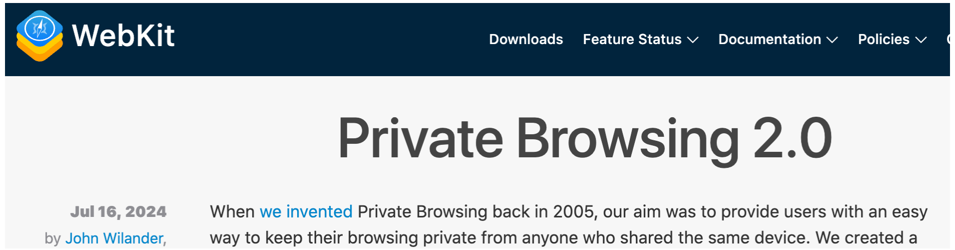
Deep perceptual hashing algorithms with hidden dual purpose: when client-side scanning does facial recognition

Year: 2023, Pages: 234-252
DOI Bookmark: 10.1109/SP46215.2023.10179310

Authors

Shubham Jain, Imperial College London
Ana-Maria Creţu, Imperial College London
Antoine Cully, Imperial College London
Yves-Alexandre de Montjoye, Imperial College London

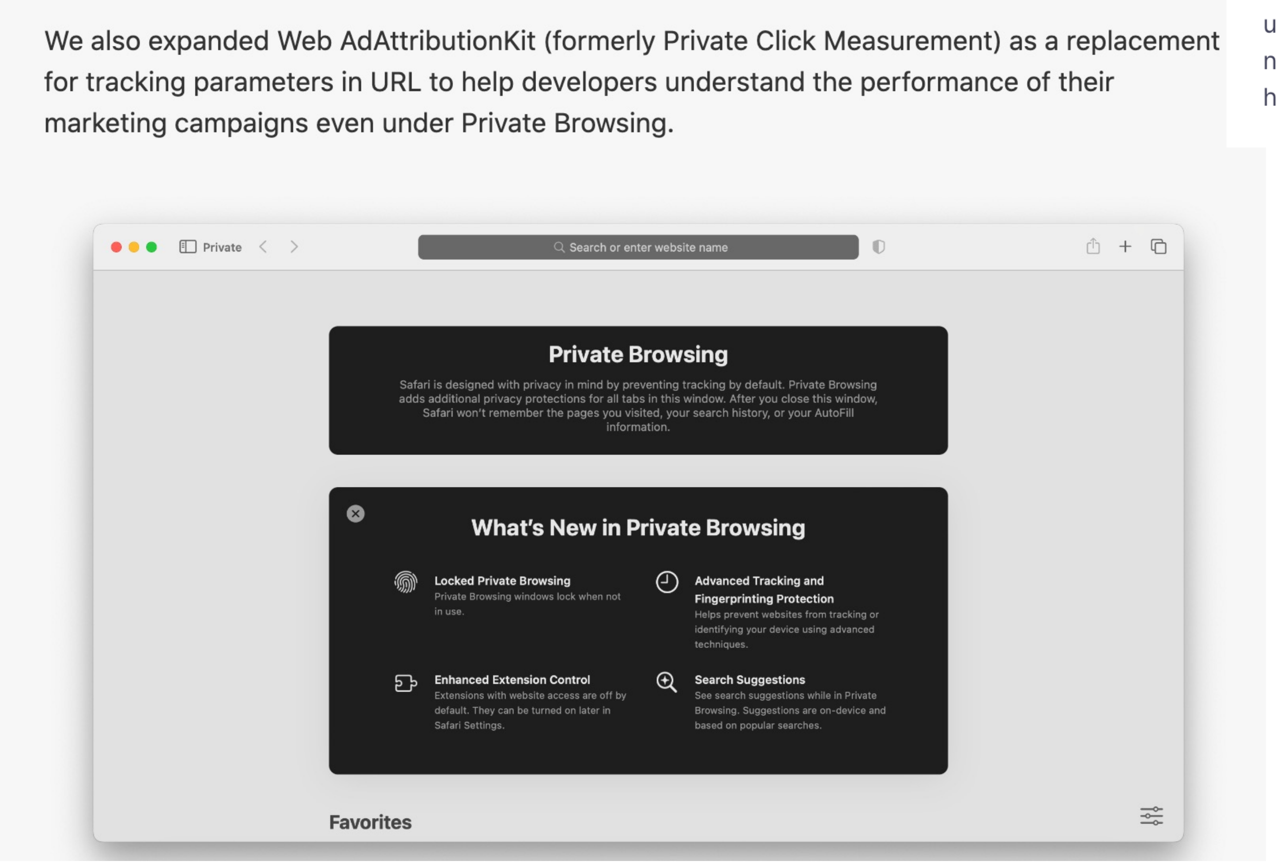
“Private” ad attribution



Private Browsing 2.0

Jul 16, 2024
by John Wilander,

When we invented Private Browsing back in 2005, our aim was to provide users with an easy way to keep their browsing private from anyone who shared the same device. We created a



We also expanded Web AdAttributionKit (formerly Private Click Measurement) as a replacement for tracking parameters in URL to help developers understand the performance of their marketing campaigns even under Private Browsing.

Private Browsing

Safari is designed with privacy in mind by preventing tracking by default. Private Browsing adds additional privacy protections for all tabs in this window. After you close this window, Safari won't remember the pages you visited, your search history, or your AutoFill information.

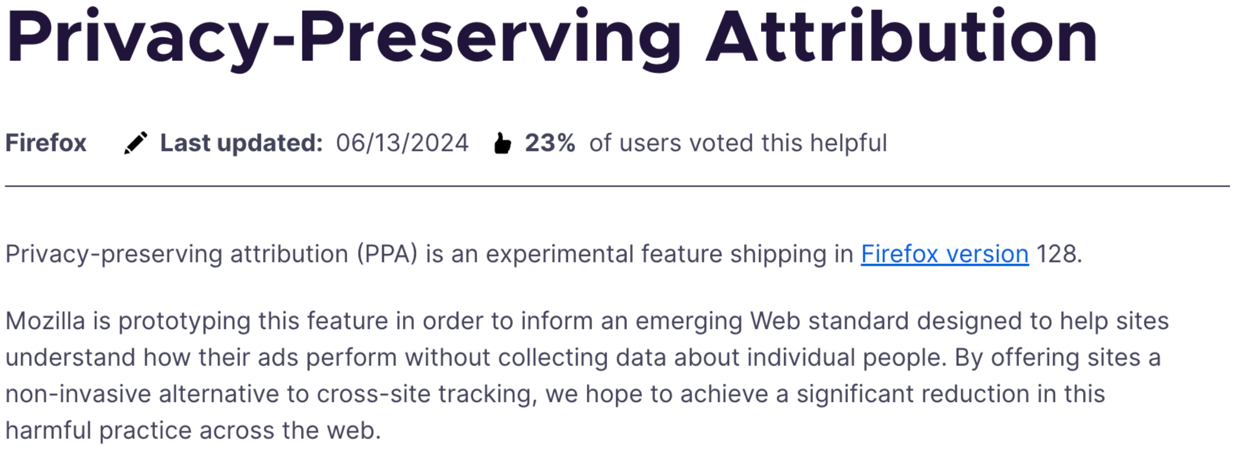
What's New in Private Browsing

Locked Private Browsing
Private Browsing windows lock when not in use.

Advanced Tracking and Fingerprinting Protection
Helps prevent websites from tracking or identifying your device using advanced techniques.

Enhanced Extension Control
Extensions with website access are off by default. They can be turned on later in Safari Settings.

Search Suggestions
See search suggestions while in Private Browsing. Suggestions are on-device and based on popular searches.



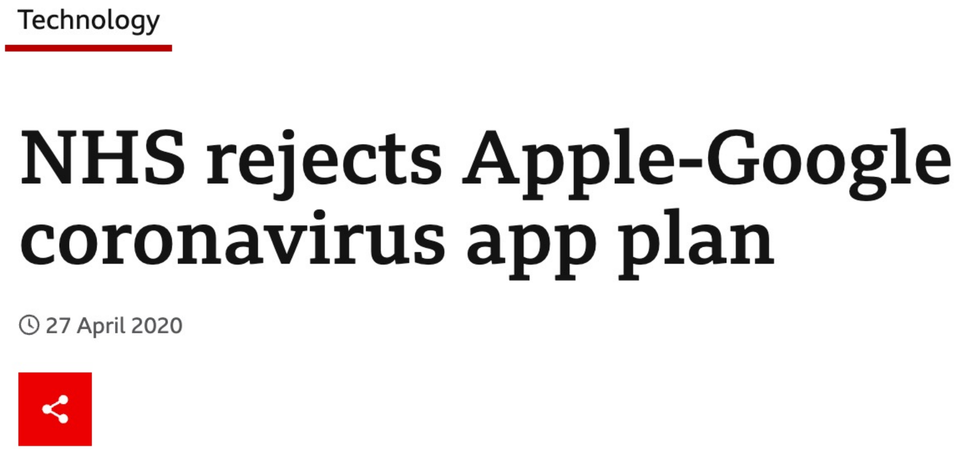
Privacy-Preserving Attribution

Firefox **Last updated:** 06/13/2024 **23%** of users voted this helpful

Privacy-preserving attribution (PPA) is an experimental feature shipping in Firefox version 128.

Mozilla is prototyping this feature in order to inform an emerging Web standard designed to help sites understand how their ads perform without collecting data about individual people. By offering sites a non-invasive alternative to cross-site tracking, we hope to achieve a significant reduction in this harmful practice across the web.

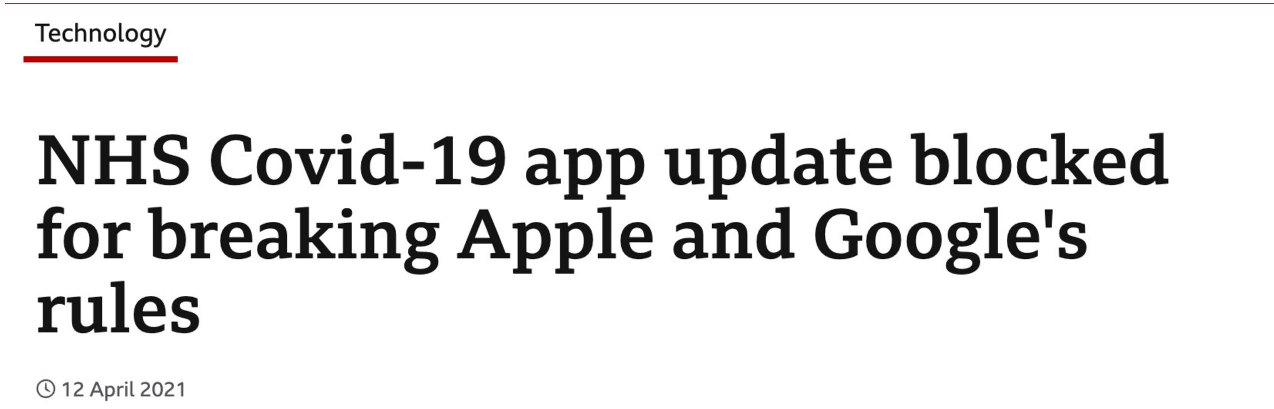
And the list goes on… contact tracing, AirTags,..



Technology

NHS rejects Apple-Google coronavirus app plan

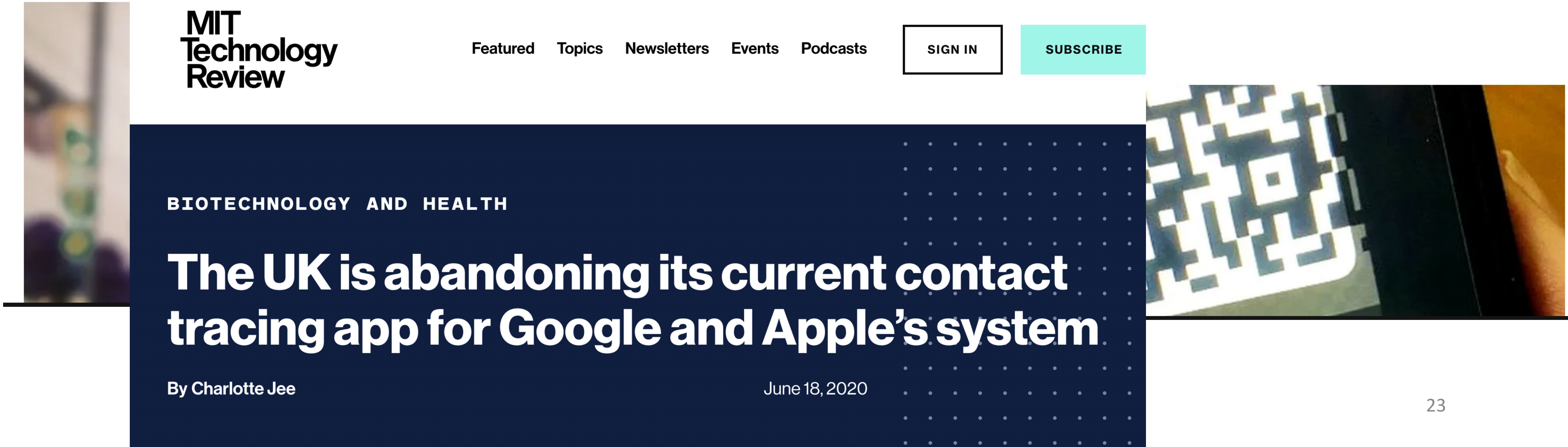
27 April 2020



Technology

NHS Covid-19 app update blocked for breaking Apple and Google's rules

12 April 2021



MIT Technology Review

BIOTECHNOLOGY AND HEALTH

The UK is abandoning its current contact tracing app for Google and Apple’s system

By Charlotte Jee

June 18, 2020

Who are we building Privacy/Security tech for?

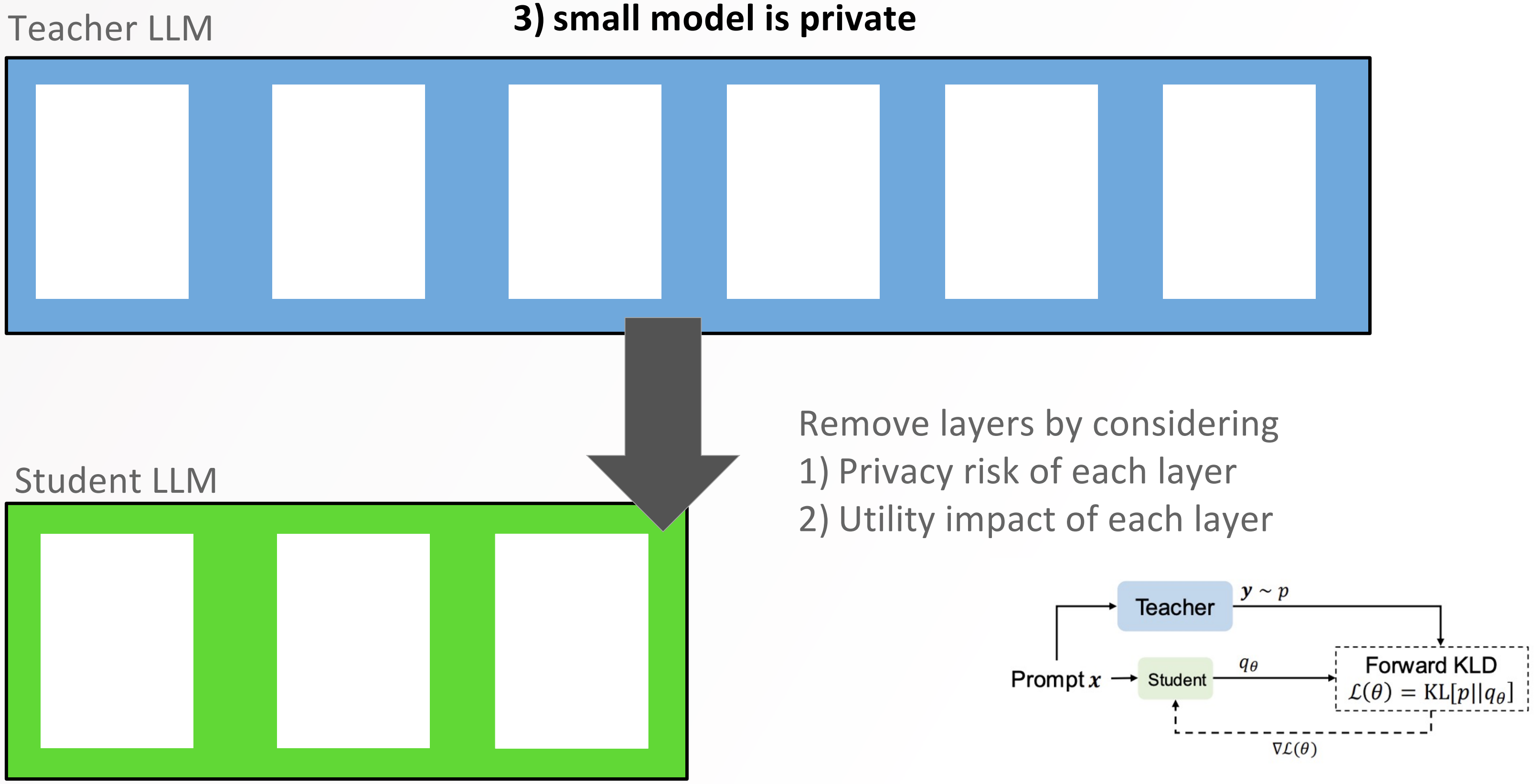
Potential solution: Use local models?

Distill knowledge to a small model such that

1) small model can be locally deployed

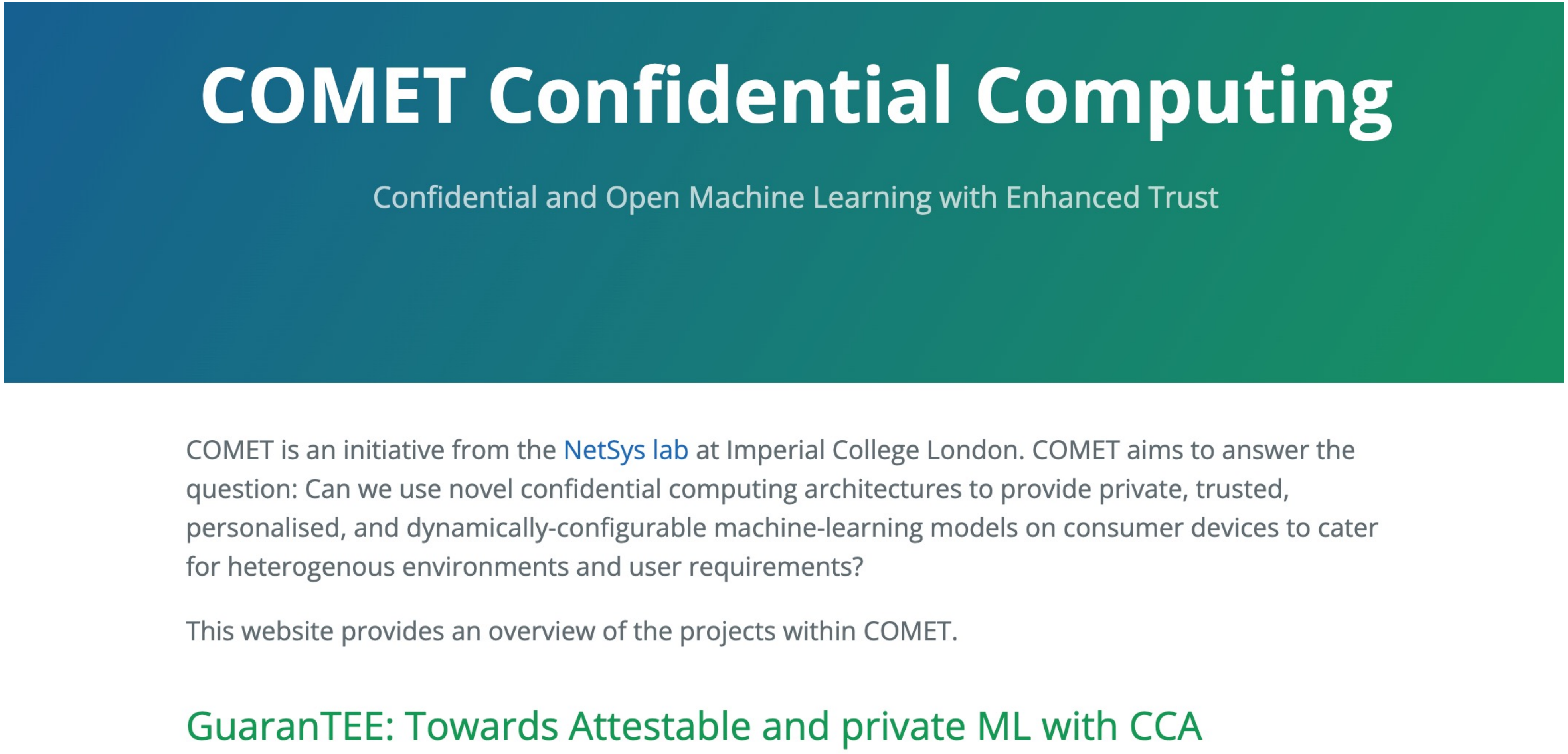
2) small model has still high utility

3) small model is private



Remove layers by considering
1) Privacy risk of each layer
2) Utility impact of each layer

Potential solution: Auditable/Confidential Computing?



COMET Confidential Computing

Confidential and Open Machine Learning with Enhanced Trust

COMET is an initiative from the NetSys lab at Imperial College London. COMET aims to answer the question: Can we use novel confidential computing architectures to provide private, trusted, personalised, and dynamically-configurable machine-learning models on consumer devices to cater for heterogenous environments and user requirements?

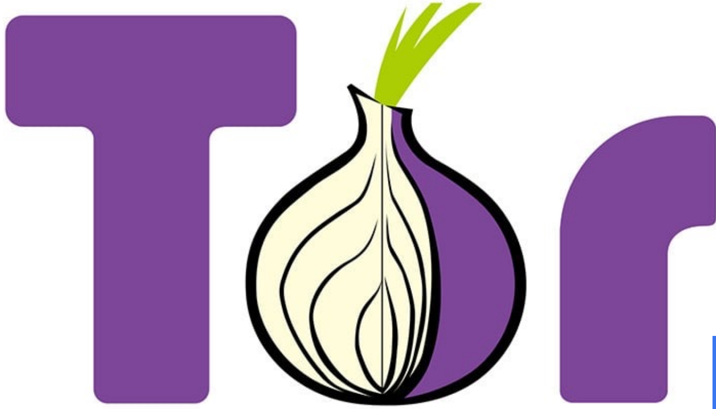
This website provides an overview of the projects within COMET.

GuaranTEE: Towards Attestable and private ML with CCA

comet-cc.github.io

Potential solutions: [FHE? SMPC? ZKP? PIR? …]

Success stories



Summary: We need to take charge!

- Personal data systems face complex challenges and exciting opportunities;
- We need to think carefully when we design and implement data collection systems;
- Trusted and auditable client-side analytics are timely enablers for privacy, security, and utility in the personal data ecosystem.

More information, software, and papers:

haddadi.github.io

IMPERIAL

brave