EPFL

Towards a safe digital society

Prof. Carmela Troncoso

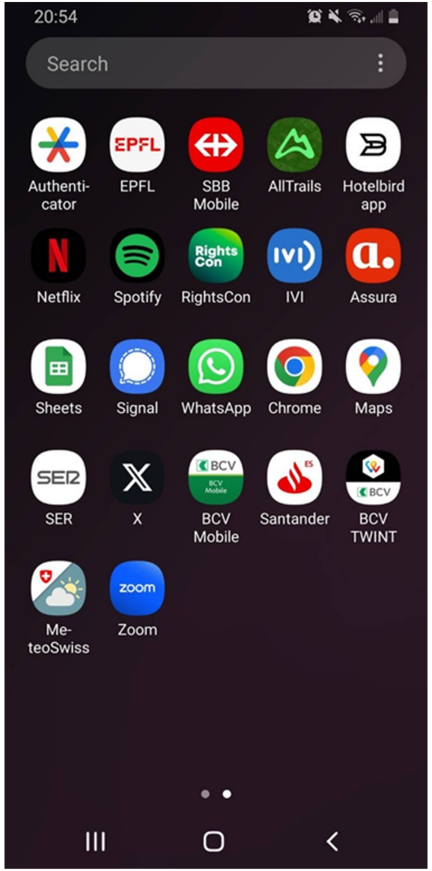
SPRING Lab

EPFL

3 May 2024

École
polytechnique
fédérale
de Lausanne

Life in our digital society...



So what if they know… I have nothing to hide

Targeted advertising

Target knows

What you buy, when you buy it, how often, …

Target can buy data about you:

Online: what webs you visit, how long, in which order, what kinds of topics you search for online, what you like, what you share,…

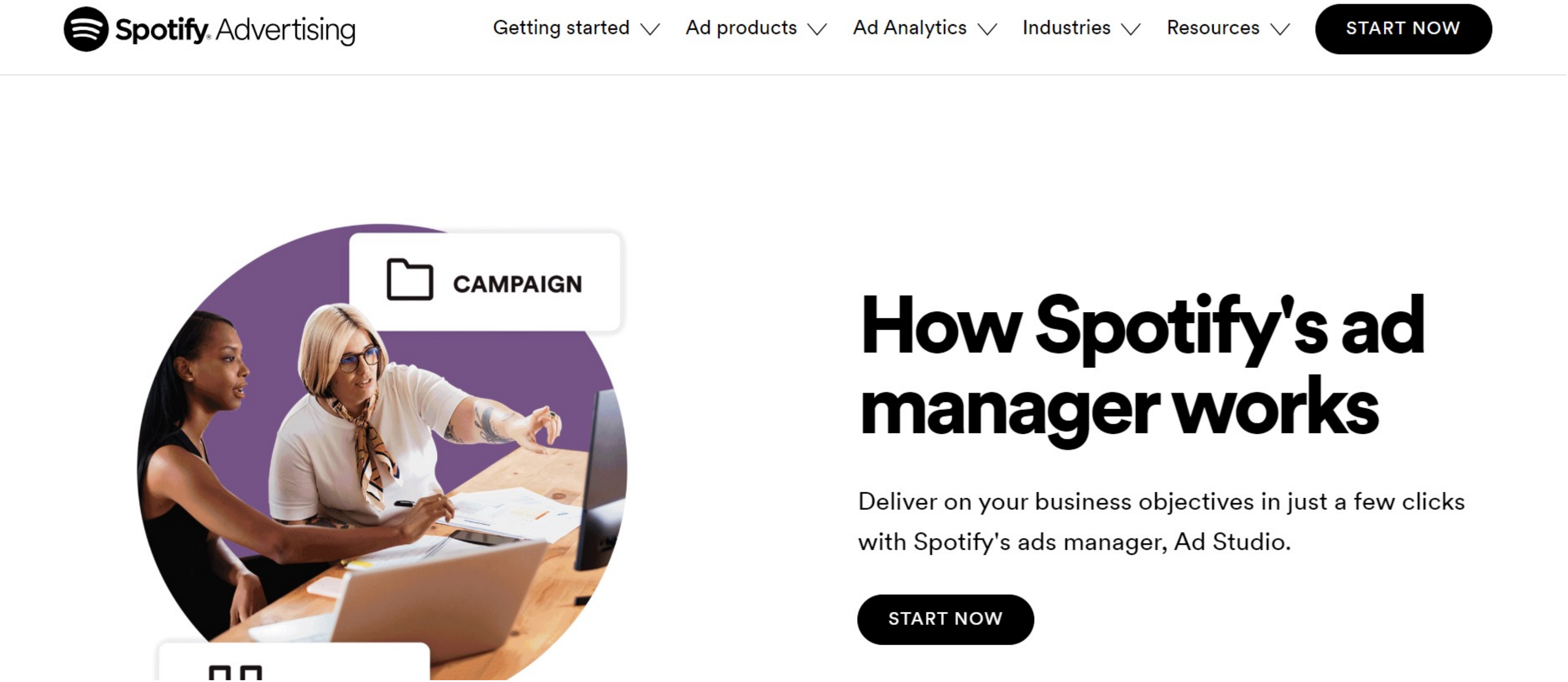
Offline: your ethnicity, job history, the magazines you read, if you've ever declared bankruptcy or got divorced, the year you bought your house, where you went to college, the number of cars you own,...

Target can use this information for

Sending customers with kids catalogs of toys before Christmas

Sending customers who buy swimsuits in April coupons for sunscreen in July and diet books in December

This may bring surprises… https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/

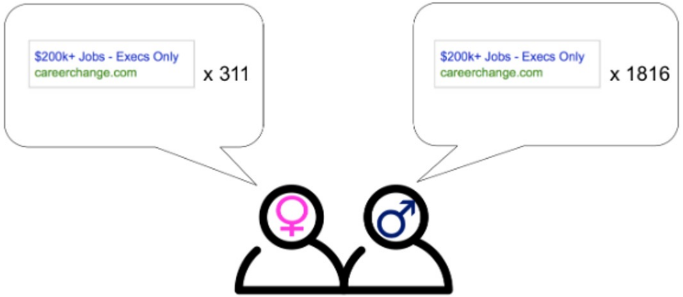


How Spotify's ad manager works

Deliver on your business objectives in just a few clicks with Spotify's ads manager, Ad Studio.

Targeted advertising

This information can also be used to discriminate

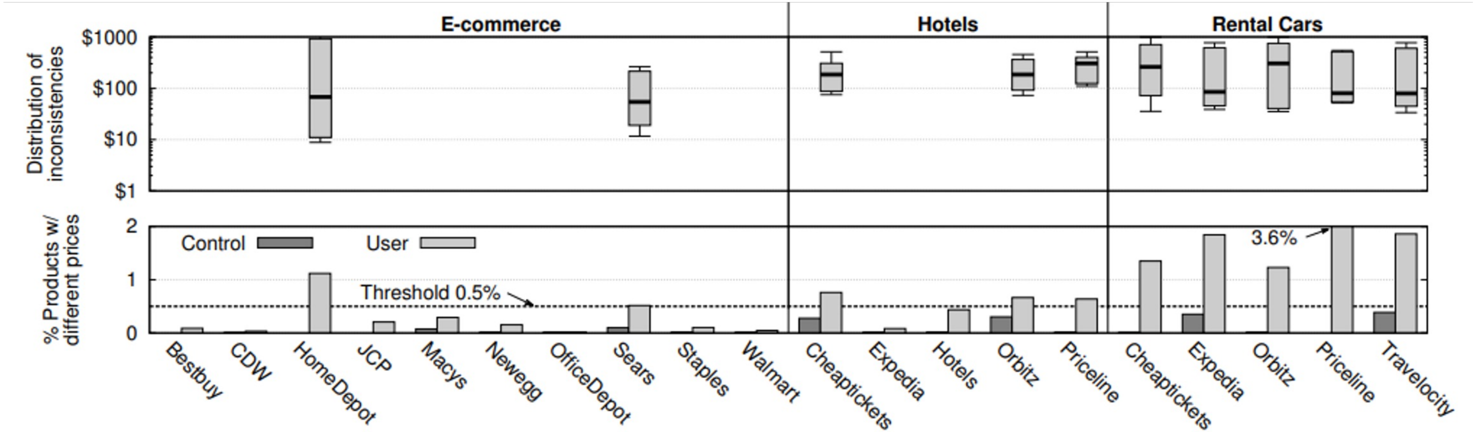


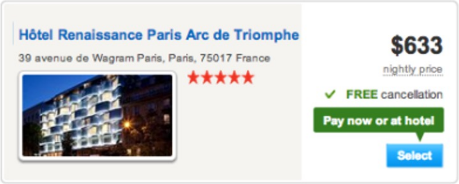
“The top two ads served to the male group was from a career coaching service called careerchange.com that promised high-paying executive level jobs. The top ad was served **1816** times to the male users, but only **311** times to the female users. Of the 500 simulated male users, **402** received the ad at least once, but only **60** female users received the same ad at least once.”

https://arxiv.org/abs/1408.6491

Targeted advertising

This information can also be used to discriminate





Hôtel Renaissance Paris Arc de Triomphe

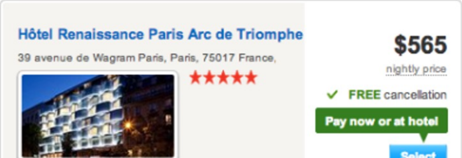
39 avenue de Wagram Paris, Paris, 75017 France

★★★★★

$633

nightly price

✓ FREE cancellation



Hôtel Renaissance Paris Arc de Triomphe

39 avenue de Wagram Paris, Paris, 75017 France,

★★★★★

$565

nightly price

✓ FREE cancellation

https://mislove.org/publications/Ecommerce-IMC.pdf

And to influence democracy….

Context: CA obtained 50M records from Facebook in 2013 through “survey” app that leaked friends’ information as well as from the user answering the survey

CA created a system that can target voters based on psychological profile

Was used to target US voters in 2016 elections and UK voters in Brexit

Why privacy is important?

Digitalization bri

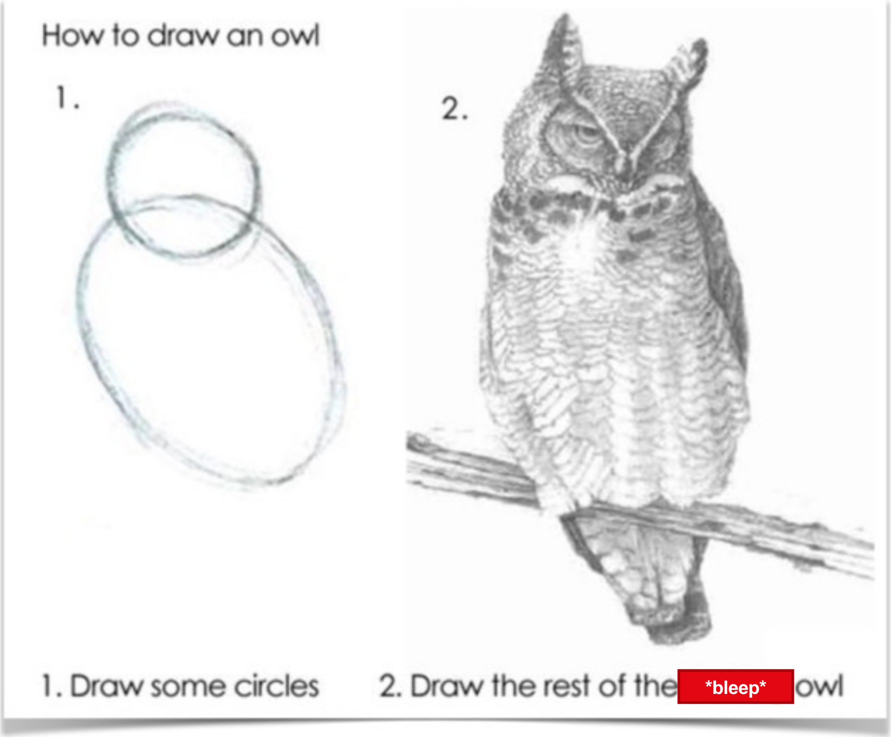
New data sou

Data collecte

Improvement

Privacy-by-desi ɔn

But how do we i



How to draw an owl

1. Draw some circles

2. Draw the rest of the *bleep* owl

Privacy is data minimization!

Let’s build systems with no data!

The least data in the system, the more privacy

Related to a legal principle – good for adoption

but, **it is not data that we minimize** (in the system as a whole)

Data are still…

in the users’ devices

encrypted at the server

distributed in servers

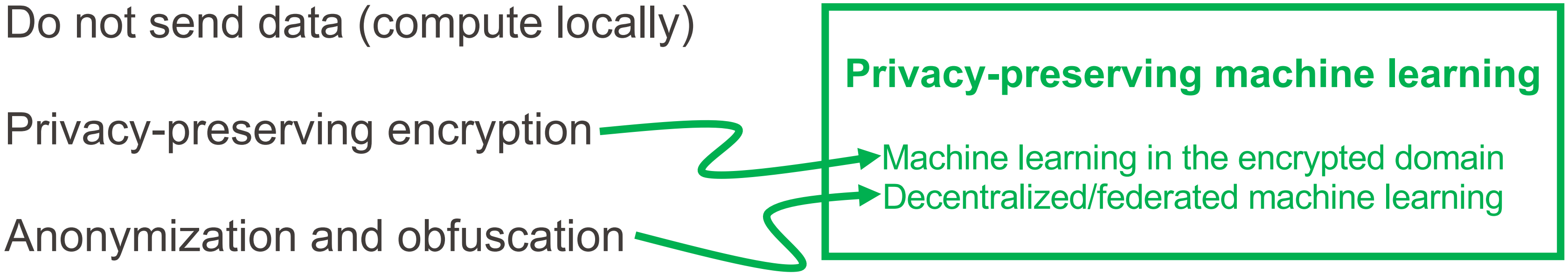
…

“data minimization” is a BAD metaphor to conceptualize designs with privacy protection

Privacy is trust minimization!

Let’s build systems in which we don’t need to trust service providers with the data!

Malicious service provider



Do not send data (compute locally)

Privacy-preserving encryption

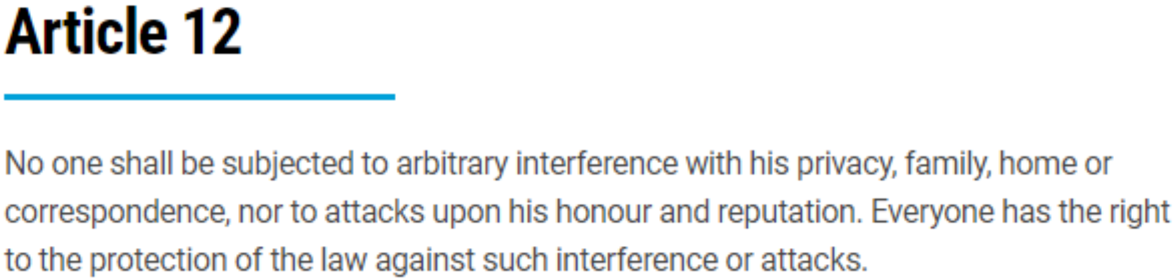
Anonymization and obfuscation

Privacy-preserving machine learning

Machine learning in the encrypted domain

Decentralized/federated machine learning

Why privacy (by design)?



Article 12

No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

https://www.un.org/en/about-us/universal-declaration-of-human-rights

Minimizing data does not guarantee no harm (or interference or attacks)
What if the minimal data still allows harms?
Or the purpose is harmful in itself? Or enables harms?

“minimizing trust” is also a BAD metaphor to conceptualize designs with privacy protection

Privacy's goal is to protect from undesired uses!

Step 1: define "**desired uses**" - the purpose of the application

Step 2: identify the **minimal data** need for this purpose

Step 3: build a system that achieves the purpose **minimizing misuse possibilities**
use Privacy Enhancing Technologies!

"purpose limitation"
is a GOOD metaphor to conceptualize designs with privacy protection

Also related to a legal principle!
good for adoption

Use case 1: protecting victims of conflict

SPRING

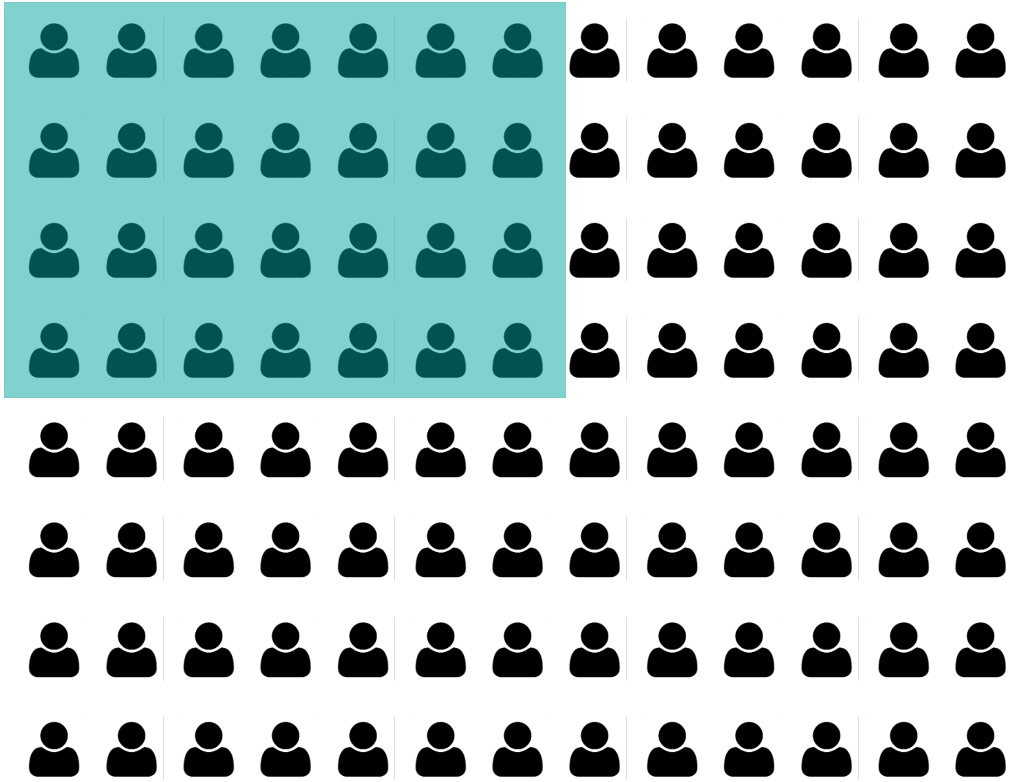
SECURITY AND PRIVACY ENGINEERING LABORATORY

COMITE INTERNATIONAL GENEVE

ICRC

Maximizing number helped people: Distribute as best as you can

Little resources…



…and lots of people to help

Humanitarian aid distribution

Traditional solution: pen and paper

Does not scale
Easy to manipulate
Hard to audit

Can we do better use digitalization to scale without introducing risks? Can biometrics help?

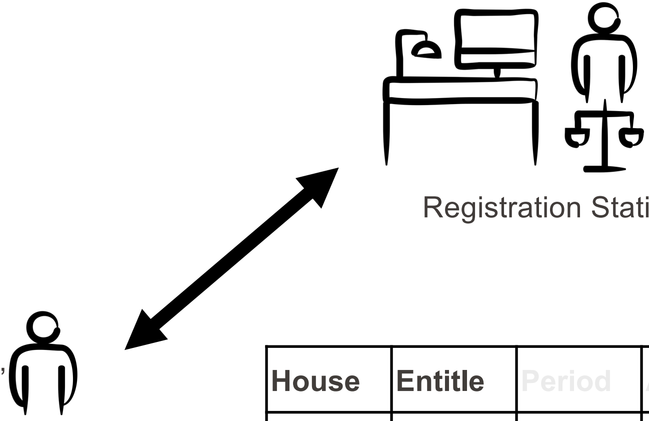
https://avarchives.icrc.org/Picture/

Humanitarian aid Requirements

Phase 1: Registration

- Registration per **household**
- **Legitimacy** check
- **Entitlement** assignment

“Yes, your **household** lives in affected area. You are entitled to **3 bags of rice & 1 baby formula**.”



Registration Station

“I am Boya from household Wang.”

Recipient

House	Entitle	Period	Auth
Wang	3+1	5	

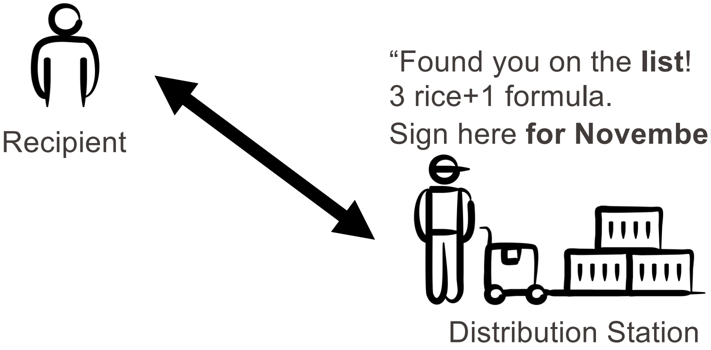
Humanitarian aid Requirements

Phase 2: Distribution

- **Legitimacy** check
- **Double-dipping** prevention
- **Periodic** Distribution

House	Entitle	Period	Auth
Wang	3+1	5	

“I request the aid for **household** Wang”



“Found you on the **list**!
3 rice+1 formula.
Sign here **for November**.”

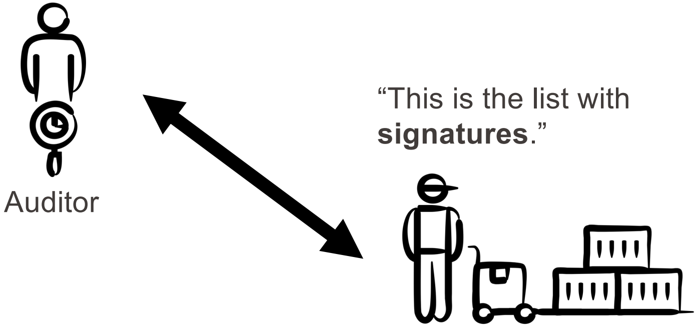
Humanitarian aid Requirements

Phase 3: Auditing

- Check distribution **proof**

House	Entitle	Period	Auth
Wang	3+1	5	

“Give me records for November. I will cross-check with warehouse.”



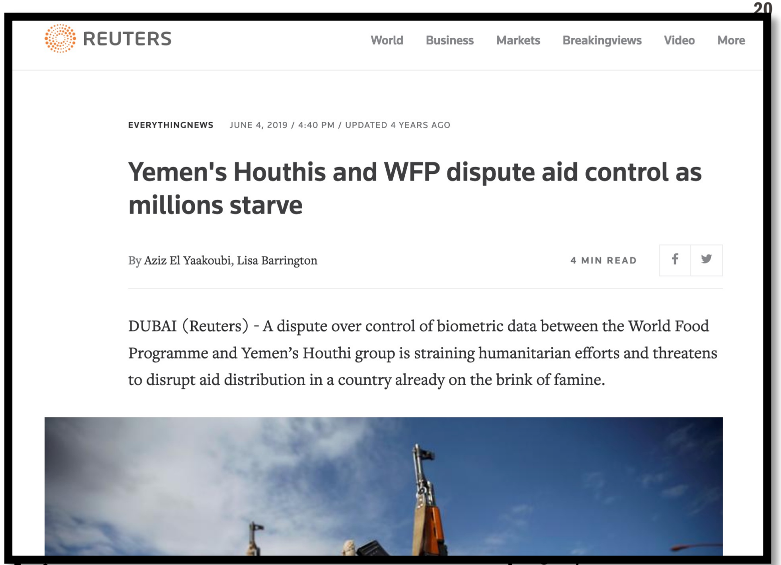
Auditor

“This is the list with **signatures**.”

Distribution Station

Straightforward digitalization

It scales but…
it does **not** prevent reuse/abuse


REUTERS

World Business Markets Breakingviews Video More

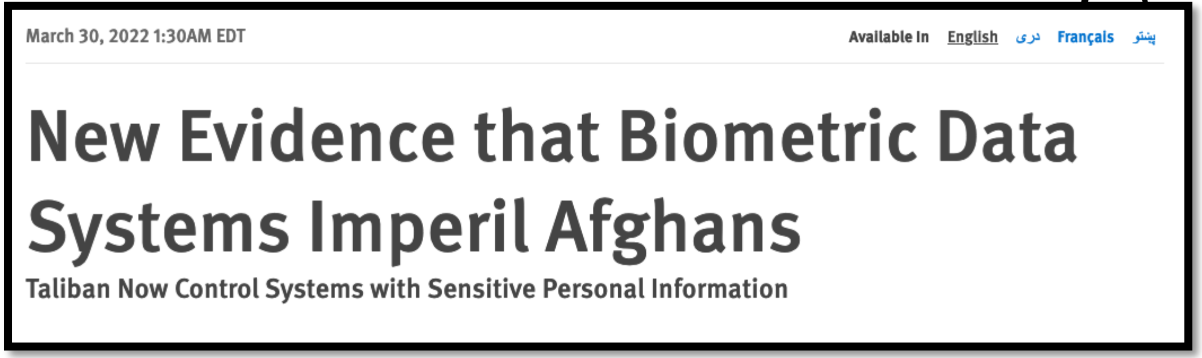
EVERYTHINGNEWS JUNE 4, 2019 / 4:40 PM / UPDATED 4 YEARS AGO

Yemen's Houthis and WFP dispute aid control as millions starve

By Aziz El Yaakoubi, Lisa Barrington

4 MIN READ

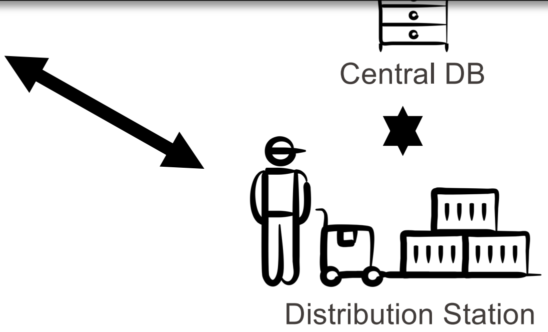
DUBAI (Reuters) - A dispute over control of biometric data between the World Food Programme and Yemen's Houthi group is straining humanitarian efforts and threatens to disrupt aid distribution in a country already on the brink of famine.


March 30, 2022 1:30AM EDT

Available In English دری Français پښتو

New Evidence that Biometric Data Systems Imperil Afghans

Taliban Now Control Systems with Sensitive Personal Information


Central DB

Distribution Station

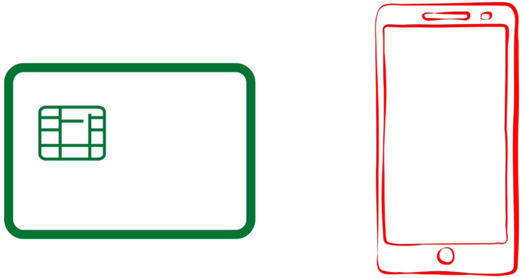
A safe solution

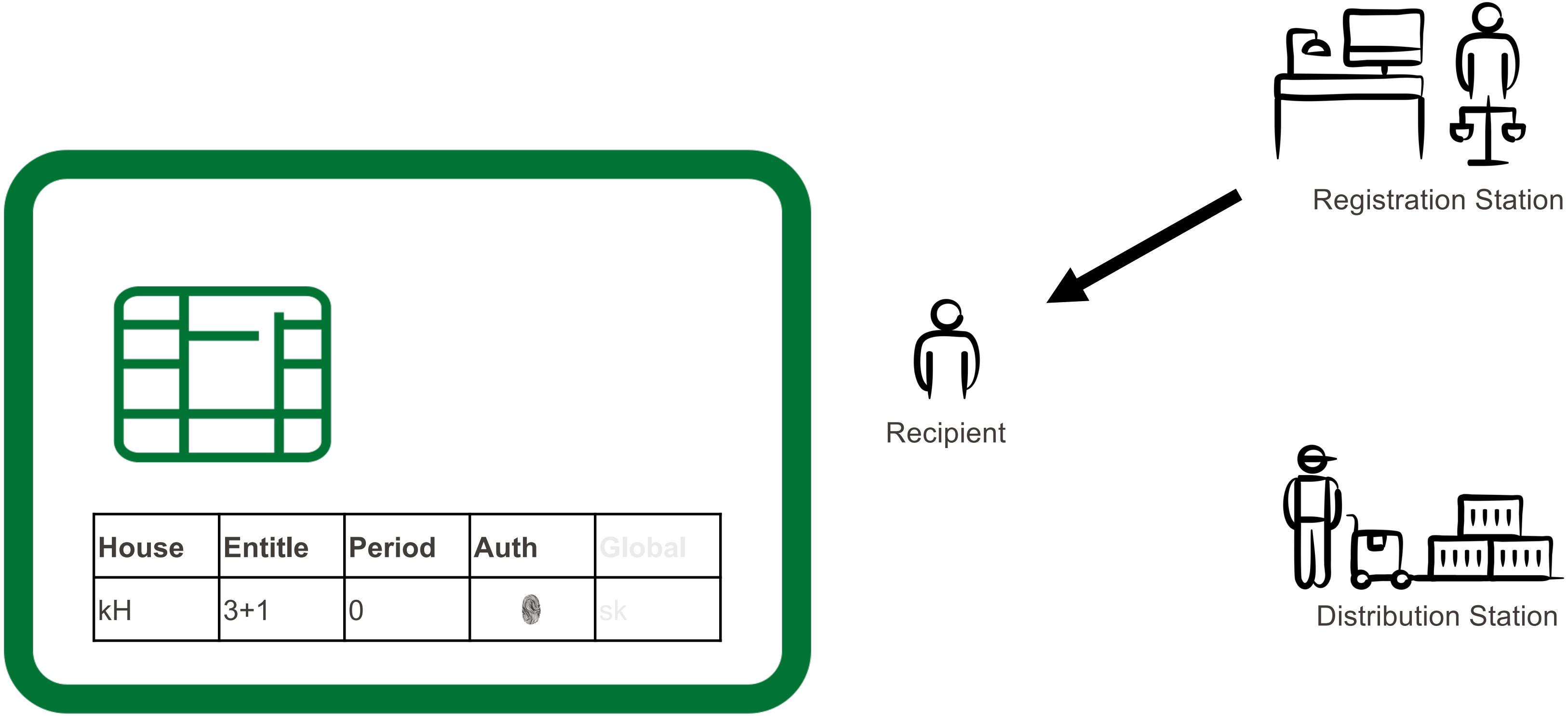
- Decentralize information in devices

  -> Legitimacy check without a database
- Unforgeable Cryptography

  -> Avoid double dipping
- Privacy-preserving cryptography

  -> Audits without recipient identification





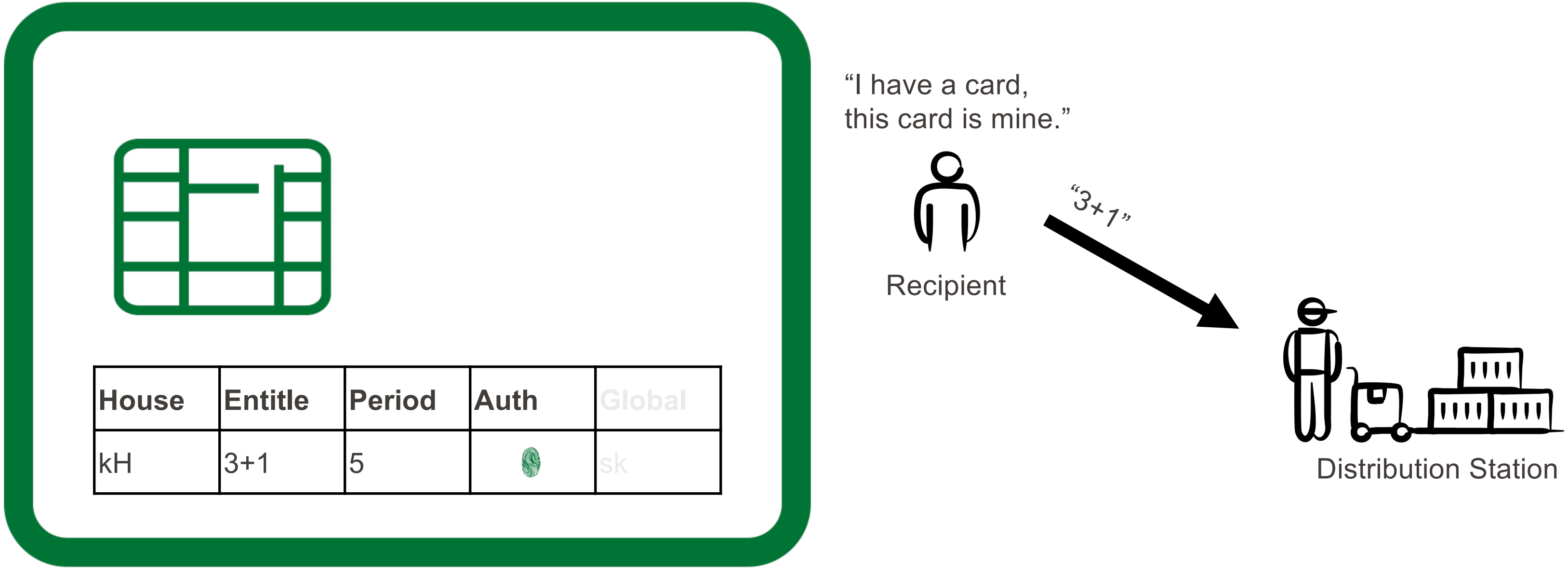
House	Entitle	Period	Auth	Global
kH	3+1	0		sk

Registration Station

Recipient

Distribution Station

- Local legitimacy check



House	Entitle	Period	Auth	Global
kH	3+1	5	●	sk

"I have a card,
this card is mine."

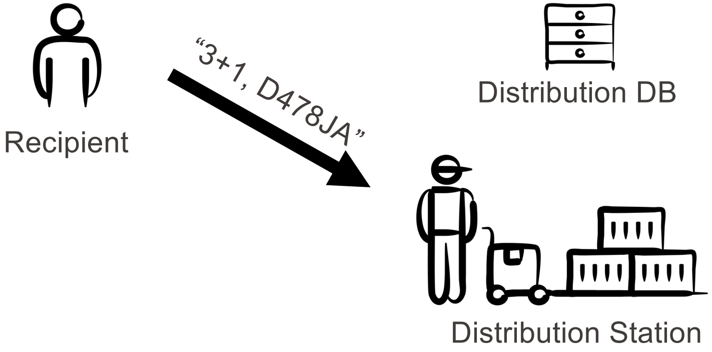
- Double dipping prevention

D478JA=PRF(kH, 5)

House	Entitle	Period	Auth	Global
kH	3+1	5		sk

“Not seen D478JA”

Ent	Period	Tag	Com	Sign
1+1	4	C3HNU0	ADBY21	BAYD24
5+2	4	2GSA8Q	BSSIA4	NDA57Y
4+3	5	NV7M91	CI79AE	34BFA1



- Privacy-preserving audit

D478JA=PRF(kH, 5)

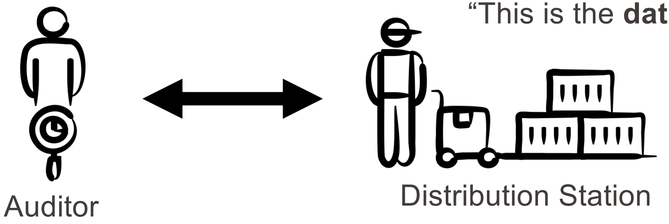
MWTX6=Commit(ent)

P9W7Z=Sign(sk, D478JA||MWTX6||5)

House	Entitle	Period	Auth	Global
kH	3+1	5	🫆	sk

Ent	Period	Tag	Com	Sign
1+1	4	C3HNU0	ADBY21	BAYD24
5+2	4	2GSA8Q	BSSIA4	NDA57Y
4+3	5	NV7M91	CI79AE	34BFA1
3+1	5	D478JA	MWTX6	P9W7Z

“This is the **database**.”



Are signatures correct? Yes: all legitimate recipients!
Duplicate tags? No: no double dipping!
Sum of entitlement = sum of commitments?
Yes: aid distributed given legitimate requests

Ent	Period	Tag	Com	Sign
1+1	4	C3HNU0	ADBY21	BAYD24
5+2	4	2GSA8Q	BSSIA4	NDA57Y
4+3	5	NV7M91	CI79AE	34BFA1
3+1	5	D478JA	MWTX6	P9W7Z

Nothing in this table can be used for anything else than intended!!!

Next steps: Pilot with the ICRC (hopefully soon)
(also working on extensions to fulfill further functionality without increasing risk)

Use case 2: helping investigative journalists

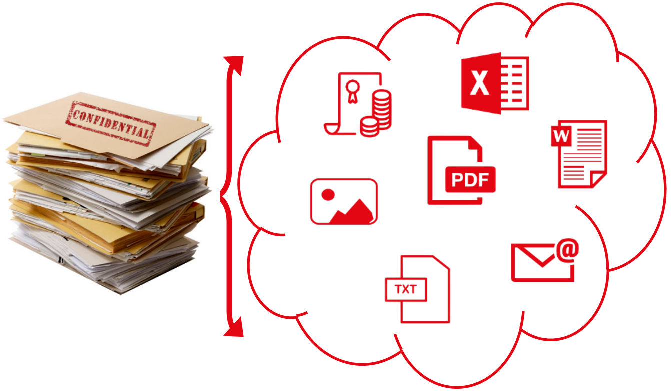
SPRING

SECURITY AND PRIVACY ENGINEERING LABORATORY

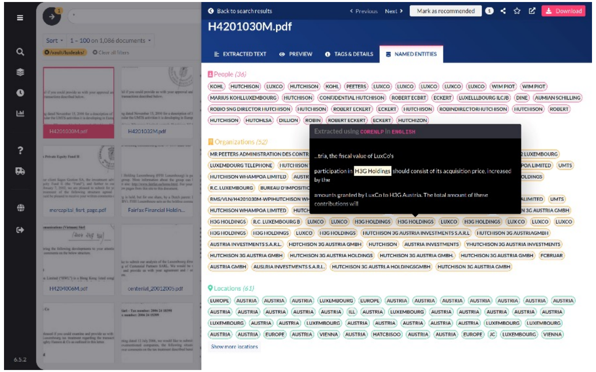
ICIJ

THE INTERNATIONAL CONSORTIUM OF INVESTIGATIVE JOURNALISTS

The problem



Leaked digital document collections are hard to search and classify



...tria, the fiscal value of LuxCo's participation in H3G Holdings should consist of its acquisition price, increased by the

amounts granted by LuxCo to H3G Austria. The total amount of these contributions will

ICIJ built a tool to **locally** index and search

Can we build safe remote search?

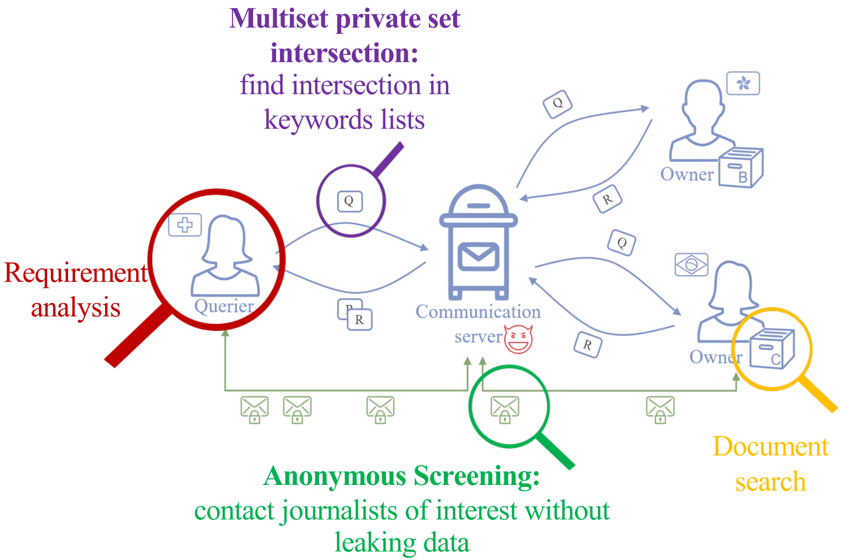
The purpose

Search: finding if others in the network have documents of interest

Contact: enabling a screening conversation before sharing

~~**Retrieval**~~

Datashare Network
End-to-end privacy engineering



Multiset private set intersection:
find intersection in keywords lists

Requirement analysis

Anonymous Screening:
contact journalists of interest without leaking data

Document search

Journalists **can search** in the network while:

- Not revealing their queries
- Not revealing their collections
- Not revealing their identities

The only learn **someone has a document of interest**

No increased risk in digitalizing!

Use case 3:
Protecting society as a whole

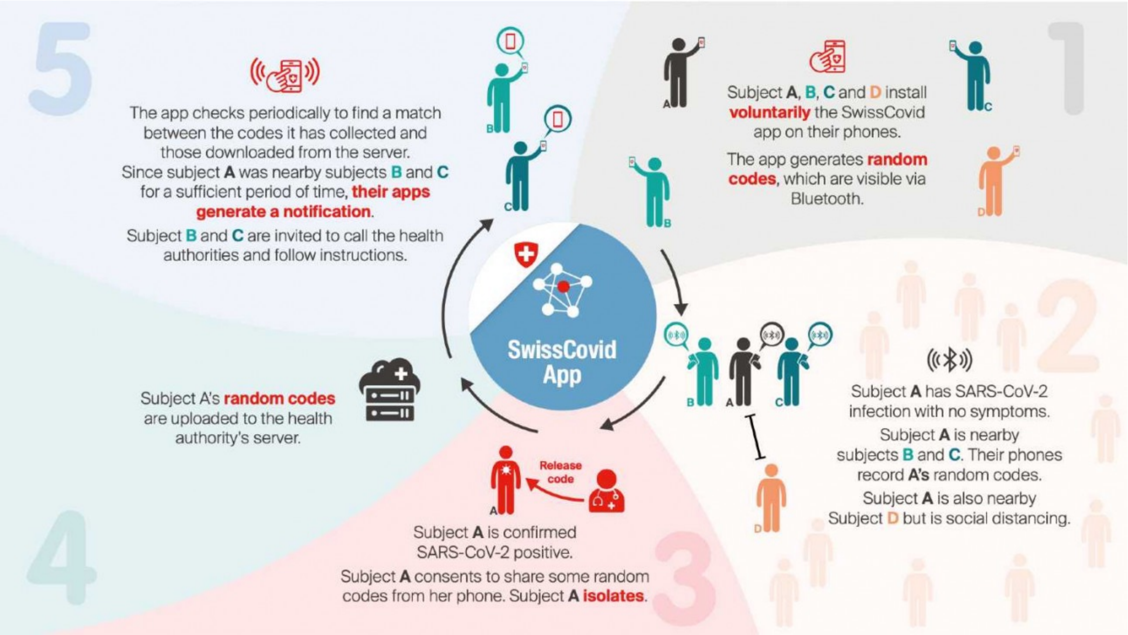
March 2020: A hard pressing problem

Covid spread too fast, contact tracing overwhelmed

A lot at stake when designing solutions

Avoid deployment of technology that can be abused in the short and long term

Decentralized privacy-preserving proximity tracing



Subject **A**, **B**, **C** and **D** install **voluntarily** the SwissCovid app on their phones.

The app generates **random codes**, which are visible via Bluetooth.

Subject **A** has SARS-CoV-2 infection with no symptoms.

Subject **A** is nearby subjects **B** and **C**. Their phones record **A's** random codes.

Subject **A** is also nearby Subject **D** but is social distancing.

Subject **A** is confirmed SARS-CoV-2 positive.

Subject **A** consents to share some random codes from her phone. Subject **A** **isolates**.

Subject A's **random codes** are uploaded to the health authority's server.

The app checks periodically to find a match between the codes it has collected and those downloaded from the server.

Since subject **A** was nearby subjects **B** and **C** for a sufficient period of time, **their apps generate a notification**.

Subject **B** and **C** are invited to call the health authorities and follow instructions.

Only information that ever leaves the phone are **random numbers broadcasted** during the contagious period (no identity, no location, no information about others)

No information available for abuse

Take-aways

Privacy **is not a goal**: it is a means to protect ourselves

Privacy engineering **must be about implementing this protection**

Privacy technologies can help minimizing harm potential
… but can also “privacy wash” harmful applications
- Client-side scanning: privacy does not limit misuse
- Privacy-preserving advertising: privacy does not limit manipulation

Limiting harm requires **limiting purpose of applications**