# On the limits of PETS when designing to prevent harm

Prof. Carmela Troncoso
SPRING Lab
MPI-SP

MAX PLANCK INSTITUTE
FOR SECURITY AND PRIVACY

# PETs are in vogue!

**Build (d)Apps with Fully Homomorphic Encryption (FHE).**

Zama is an open source cryptography company building state-of-the-art FHE solutions for blockchain and AI.

**Home » Zama Raises $73M in Series A Led by Multicoin Capital and Protocol Labs to Commercialize Fully Homomorphic Encryption**

Apple uses local differential privacy to help protect the privacy of user activity in a given time period, while still gaining insight that improves the intelligence and usability of such features as:

- QuickType suggestions
- Emoji suggestions
- Lookup Hints
- Safari Energy Draining Domains
- Safari Autoplay Intent Detection (macOS High Sierra)
- Safari Crashing Domains (iOS 11)
- Health Type Usage (iOS 10.2)


Google for Developers

Search all articles... Search

As new digital platforms and services emerge, the challenge of keeping users' information safe online is growing more complex – novel technologies require novel privacy solutions. At Google, we continue to invest in privacy-enhancing technologies (PETs), a family of cutting-edge tools that help solve the critical task of data processing by providing people guarantees that their personal information is kept private and secure.

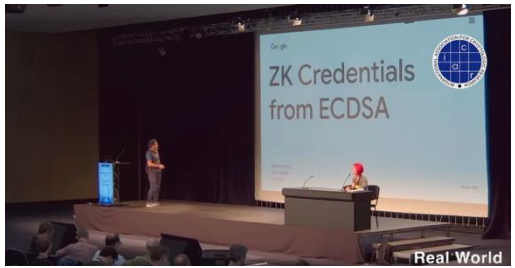
Over the past decade, we've integrated PETs throughout our product suite, used them to help tackle societal challenges and made many of our own freely available to developers and researchers around the world via open source projects.

Privacy Sandbox

Attribution Reporting: generating summary reports

On this page
Implementation status
What is an Attribution Reporting summary report?
Why do we need summary reports?
How is user data captured and aggregated?
Generate reports with the Aggregation Service
What information is captured?
How is browser data captured before aggregation?
Filtering IDs
...

Measure ad conversions aggregated across users, without revealing individual data. Formerly known as aggregate reports.




Cloudflare Docs

On this page › Overview

## Privacy Pass

Copy Page

Privacy Pass ↗ specifies an extensible protocol for creating and redeeming anonymous and transferable tokens. Its specification is maintained by the IETF. Cloudflare provides "Silk - Privacy Pass Client". This is a Chrome and Firefox browser extension used for research, which provides a better visitor experience for Cloudflare-protected websites. Privacy Pass is especially helpful for visitors from shared networks, VPNs, and Tor that tend to have poorer IP reputations.

# All these advances assume privacy = data confidentiality

The least data is available, the more privacy -- let's build systems in which we don't need to trust service providers with data

Throwing crypto/differential privacy/TEE at the problem to enable private processing is a neat solution

From a confidentiality perspective… the right to privacy is guaranteed!

# But is it?

## 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

**Does confidentiality prevent arbitrary interference?**


100K users installed CA Facebook App

enabled **COLLECTING PERSONAL DATA** of 87+ million public profile, page likes, birthday and city

creation of **PROFILES** of the subjects of the data

**TARGETED ADVERTISEMENTS**

influenced the 2016 US elections

influenced the Brexit vote


10

enable

publ

lion

creation of **PROFILES** of the subjects of the data



016 US elections

influenced the Brexit vote

# And if this happened nowadays…?

With all the new shiny PETs being developed…

- Facebook could offer an API to obtain differentially-private data of their users to train the model
- Facebook and CA could have run a cutting-edge MPC model to train CAs models
- CA could offer Facebook millions to input data to an FHE model

All these options guarantee data confidentiality! Privacy is guaranteed!

# Would our rights be protected?

## 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

A model trained in a privacy-preserving way enables the same inferences as a model trained on the data (or so these PETs developers promise…)

Confidential computing would not have prevented CA from influencing democratic processes

# European Digital Identity (EUDI)

**Core Idea**: a digital ID that improves trust in digital services

Actually… a wallet of trustworthy attributes (including identity)

**Privacy**

**Alarm!**

Attributes all over the place!
Linkability! Profiling!
Cambridge Analytica all over??

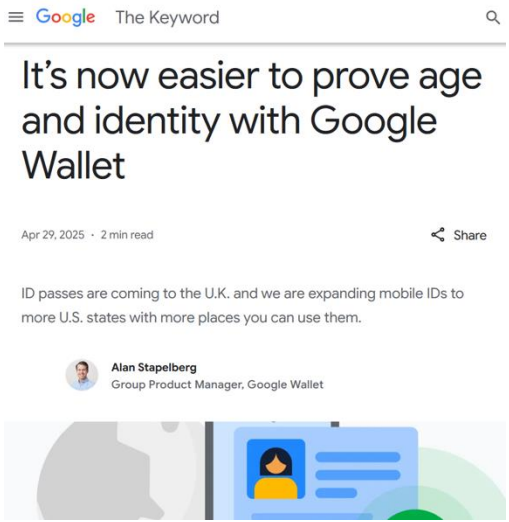
# Crypto to the rescue!

The crypto community knows how to show this problem!

Decades of research on anonymous credentials / attribute-based credentials

*(Modulo reality is less simple than research papers – see Anja Lehmann RWC25 talk)*

And even the giants have finally decided that implementing credentials is worthy for business



# Would our rights be protected?

## 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

**Problem 1:** what attributes are revealed or how is not a crypto problem

aka, *function creep* is likely to happen (especially given that the EC has made Relying party certificates that establish who can ask for what optional)

**AND** interference based on attributes is still possible

# Would our rights be protected?

## Problem 2: Digital identity problems go well beyond privacy.

**Aadhaar Linked To Half The Reported Starvation Deaths Since 2015, Say Researchers**

The Supreme Court declared today that the Aadhaar is constitutionally valid. This week also marks a year since 11-year-old Santoshi Kumari starved to death in Jharkhand because her family's ration card was not linked to Aadhaar.

Newsroom

Sep 25, 2018, 09:34 PM EDT
**Updated** Sep 26, 2018

### Real Choice & Data Safety

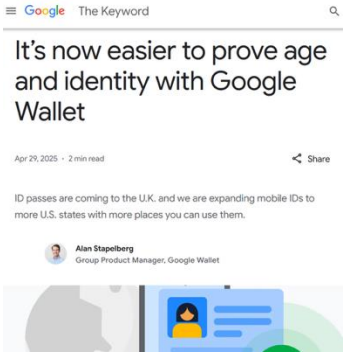
First of all, it is essential that every potential user has a real choice about using or not using the new digital ID. Without strong non-discrimination protections in the law, those who don't want or aren't able to use the new digital ID will be left out. **Nobody should be at a disadvantage for example just because they don't have a smartphone**. Those who choose to use the new ID, on the other hand, should be able to rest assured that nobody spies on which services they use and whatever information they share with this digital ID – neither governments nor private enterprises.

https://edri.org/our-work/european-digital-identity-a-potential-game-changer/

**Improving privacy might privacy-wash undesired consequences of (EU) Digital Identity**

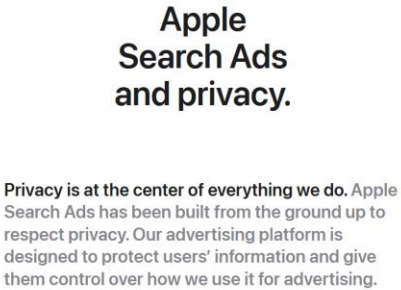
# Would our rights be protected?

## Problem 3: Privacy brings power (of decision)




**What can be proved?**


**What categories can be used for targetting?**




**Interference is still possible and when and how it can happen, decided by the few implementing the technologies**

# We must change the design paradigm!

## 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

Privacy is not about protecting data
It is about protecting people from the consequences of processing the data

**Solution:** limit the purposes for which data can be processed
**From data minimization to purpose limiting**

# Two examples

## Decentralized Privacy-Preserving Proximity Tracing

Version: 25 May 2020.

Contact the first author for the latest version.

**EPFL**: Prof. Carmela Troncoso, Prof. Mathias Payer, Prof. Jean-Pierre Hubaux, Prof. Marcel Salathé, Prof. James Larus, Prof. Edouard Bugnion, Dr. Wouter Lueks, Theresa Stadler, Dr. Apostolos Pyrgelis, Dr. Daniele Antonioli, Ludovic Barman, Sylvain Chatel

**ETHZ**: Prof. Kenneth Paterson, Prof. Srdjan Čapkun, Prof. David Basin, Dr. Jan Beutel, Dr. Dennis Jackson, Dr. Marc Roeschlin, Patrick Leu

**KU Leuven**: Prof. Bart Preneel, Prof. Nigel Smart, Dr. Aysajan Abidin

**TU Delft**: Prof. Seda Gürses

**University College London**: Dr. Michael Veale

## Not Yet Another Digital ID: Privacy-Preserving Humanitarian Aid Distribution

Boya Wang*, Wouter Lueks†, Justinas Sukaitis‡, Vincent Graf Narbel‡, Carmela Troncoso*
*SPRING Lab, EPFL, Lausanne, Switzerland
{boya.wang,carmela.troncoso}@epfl.ch
†CISPA Helmholtz Center for Information Security, Saarbrücken, Germany
lueks@cispa.de
‡International Committee of the Red Cross, Geneva, Switzerland
{jsukaitis,vgraf}@icrc.org

***Abstract*—Humanitarian aid-distribution programs help bring physical goods to people in need. Traditional paper-based solutions to support aid distribution do not scale to large populations and are hard to secure. Existing digital solutions solve these issues, at the cost of collecting large amount of personal information. This lack of privacy can endanger recipients' safety and harm their dignity. In collaboration with the International Committee of the Red Cross, we build a safe digital aid-distribution system. We first systematize the**

requires a deep understanding of the humanitarian context. We partner with the ICRC to learn the requirements and constraints associated with distributing aid in emergencies. Our interactions reveal the following challenges:

1) *Secure household-oriented aid.* Aid-distribution systems must permit aid allocation per household (i.e., a domestic unit of several members sharing meals and income), yet they must ensure that households can only request aid once per distribution round (e.g., per month).

# A blast from the past

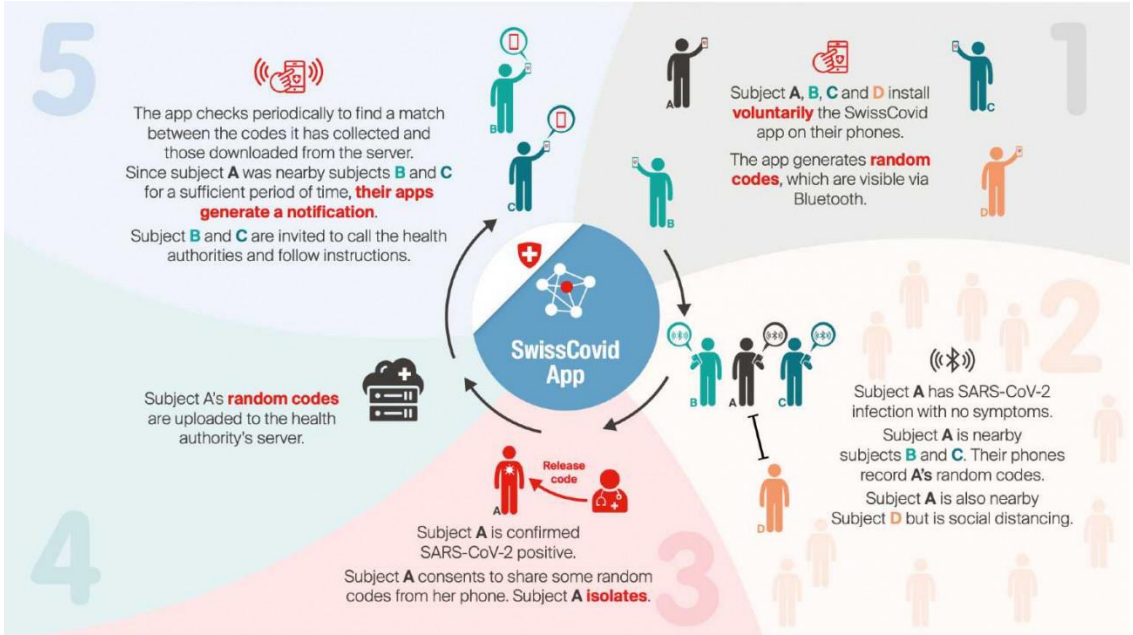
## March 2020: A hard pressing problem

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

# Privacy was the means




**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

(and easy dismantling)

# Helping victims of conflict


# Humanitarian aid distribution

## Traditional solution: pen and paper

Does not scale
Easy to manipulate
Hard to audit

***Can we do better digitalizing?***
***Can we scale and be secure without creating new risks?***

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

# Humanitarian aid requirements

## Frequent meetings with DPO office. Workshop with ECOSEC workers

### Registration

- Registration per **household** & **entitlement assigment**

"Yor **household** lives in affected area. You are entitled to **3 bags of rice & 1 baby formula**."

### Distribution

- **Periodic** distribution to **legitimate** recipients avoiding **double dipping**

| House | Entitle | Period | Auth |
|---|---|---|---|
| Wang | 3+1 | 5 | |

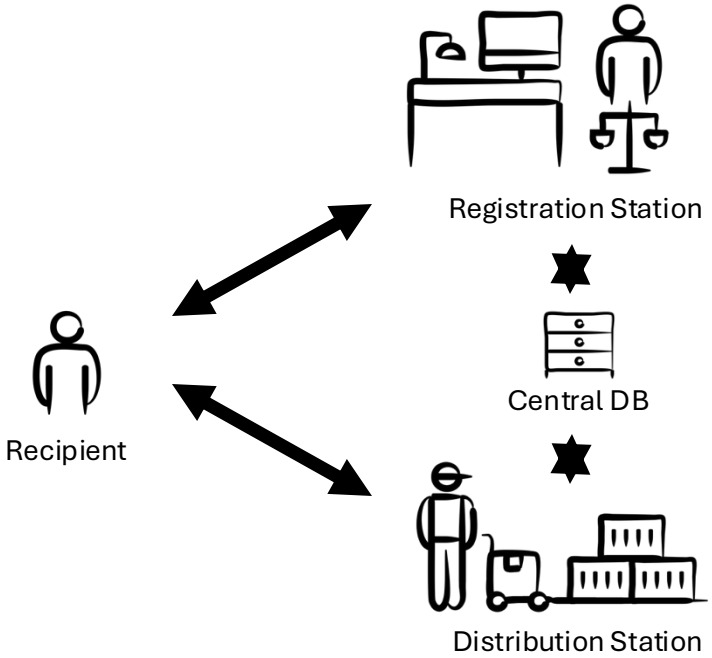
### Audits

- Provide **proof** of distribution to check against warehouse

# Straightforward digitalization

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

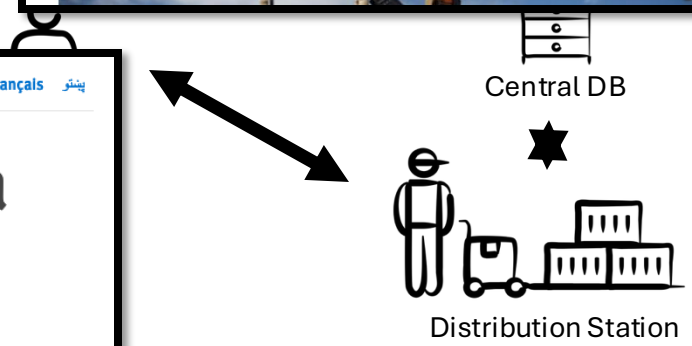
| House | Entitle | Period | Auth |
|---|---|---|---|
| Wang | 3+1 | 5 | |



# Straightforward digitalization

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





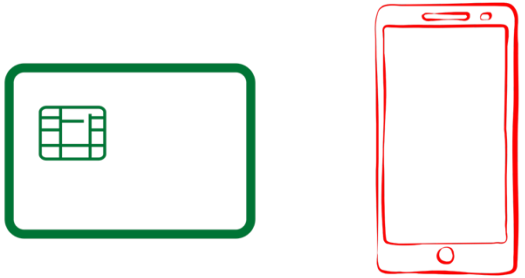
# Our solution

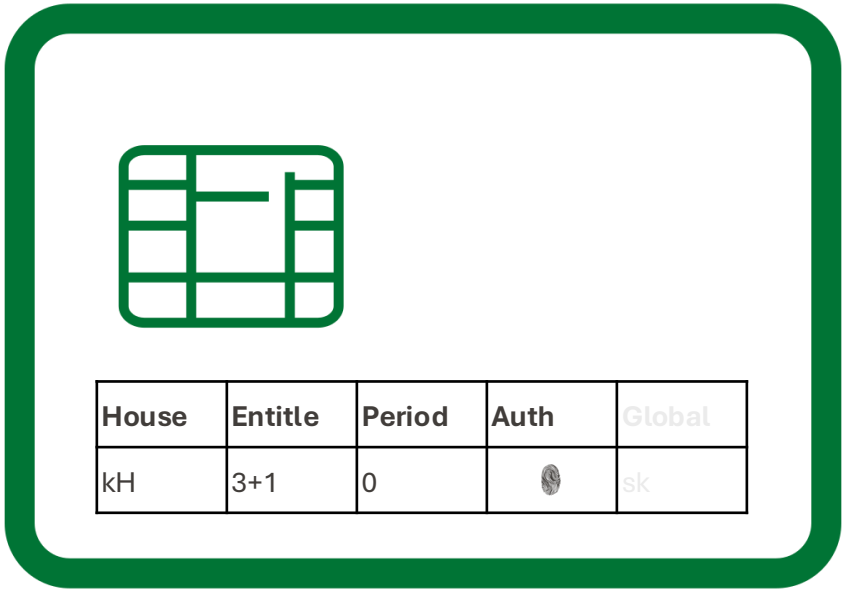
- Decentralize information in devices

  -> Legitimacy check without a database
- Unforgeable Cryptography

  -> Avoid double dipping
- Privacy-preserving cryptography

  -> Audits without recipient identification



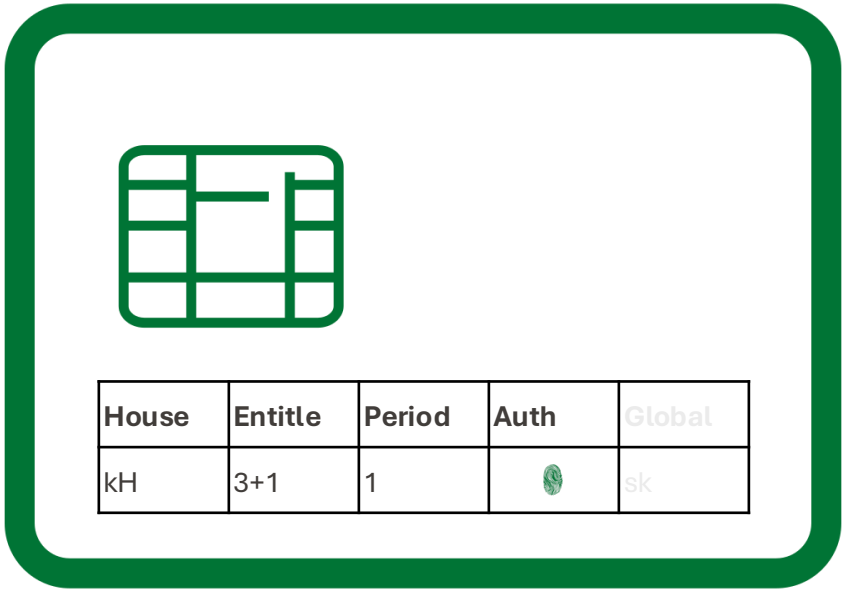



Registration Station

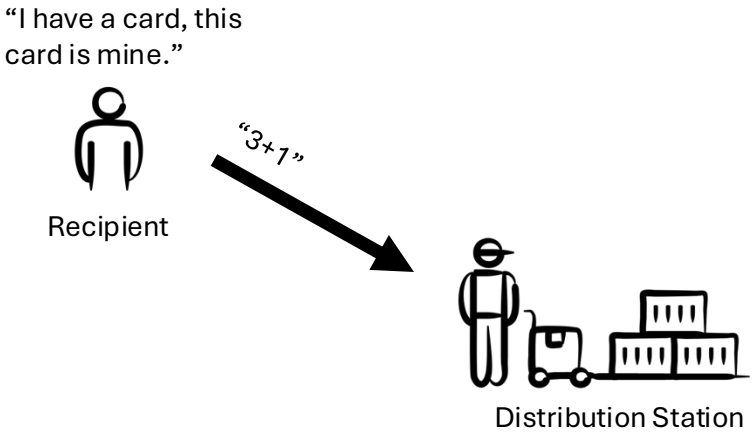
Recipient

Distribution Station

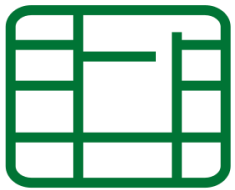
- Local legitimacy check



| House | Entitle | Period | Auth | Global |
|---|---|---|---|---|
| kH | 3+1 | 1 | | sk |



- Double dipping prevention

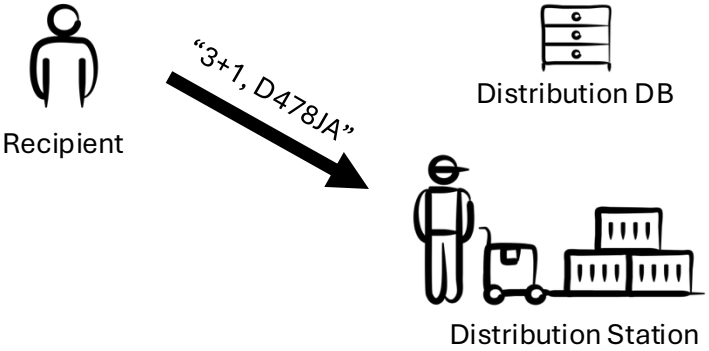


D478JA=PRF(kH, 5)

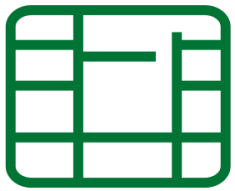
| House | Entitle | Period | Auth | Global |
|---|---|---|---|---|
| kH | 3+1 | 5 | | sk |

“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



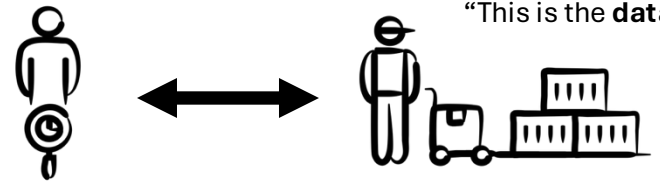
D478JA=PRF(kH, 5)

**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 |




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!!!**

# A process towards implementing Art 12 UDHR

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

NON-TRIVIAL!

2-6 months of work, read internal documentation, conversations different teams and weekly interactions with our stakeholders to never miss the point

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

Step 3: build a system that achieves the purpose **minimizing misuse possibilities** using Privacy Enhancing Technologies! (Privacy is a means! Not an end)

# A perk of purpose limitation

Purpose defines the **fundamental leakage** of a system, and the fundamental leakage determines the **inherent harm**

- Imagine the perfect system that outputs the minimal amount of bits: what inferences can you do with it? What harms do they cause?

*Corollary 1*: if those harms are not acceptable, don't deploy the system; PETs won't help you

*Corollary 2*: if the harms are acceptable, privacy-preserving design might be easier: no need to protect the fundamental leakage

# A consequence of purpose limitation

Implementing privacy brings power

(Let’s not have privacy implemented by the few!)

# And when purpose can’t be pre-defined?

We need new (privacy?) definitions to talk about purpose! To identify systems that do not technically limit interference (and thus harms)

Can we capture the risk of function creep in a definition?

Can we quantify function creepiness?

Two current very relevant cases:

EUDI: can we quantify which attributes lead to what harms?

Designated-Commitment TLS: can we quantify the privacy loss when statements are proved?

# Takeaways

Doing privacy engineering as data confidentiality risks forsaking fundamental rights

Instead, PETs should help enabiling purpose limitation

We need means to evaluate purpose specification - not (only) a policy matter

We must acknowledge the fundamental leakage and associated harms of systems as part of their issues


https://www.dagstuhl.de/25112





# Interested on discussing/working on this?

Let’s talk today!

Visit the SPRING group at MPI!

I am looking for crypto-jedi post-docs

MPI-SP has open tenure-track positions: apply!