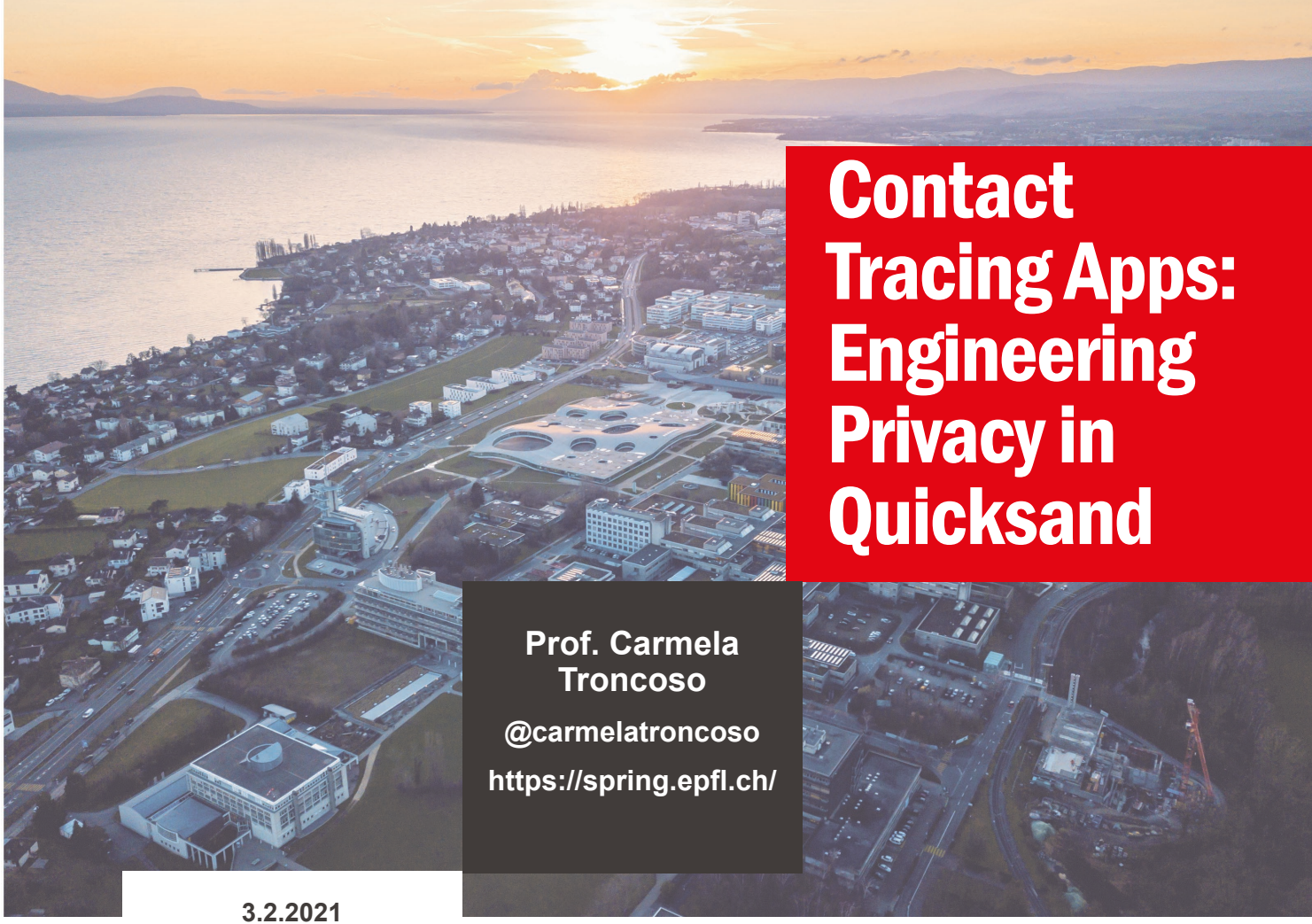
Contact Tracing Apps: Engineering Privacy in Quicksand



Prof. Carmela Troncoso

@carmelatroncoso

https://spring.epfl.ch/

3.2.2021

École polytechnique fédérale de Lausanne

A collaborative (continued) sprint

Maintenance and support

March 2020 - **Start**

April 2020 – **GAEN is announced**

May 2020 – **Final version DP3T**

June 2020 – **Pilot SwissCovid (& other EU apps)**

July 2020 – **SwissCovid launch**

August/September 2020 – **Towards international interoperability**

September/November 2020 – **Presence tracing**

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

CISPA: Prof. Cas Cremers, Prof. Michael Backes, Dr. Nils Ole Tippenhauer

University of Oxford: Dr. Reuben Binns

University of Torino / ISI Foundation: Prof. Ciro Cattuto

Aix Marseille Univ, Université de Toulon, CNRS, CPT: Dr. Alain Barrat

IMDEA Software Institute: Prof. Dario Fiore

INESC TEC: Prof. Manuel Barbosa (FCUP), Prof. Rui Oliveira (UMinho), Prof. José Pereira (UMinho)

A collaborative (continued) ~~sprint~~ Marathon

Maintenance and support

March 2020 - **Start**

April 2020 – **GAEN is announced**

May 2020 – **Final version DP3T**

June 2020 – **Pilot SwissCovid (& other EU apps)**

July 2020 – **SwissCovid launch**

August/September 2020 – **Towards international interoperability**

September/November 2020 – **Presence tracing**

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

CISPA: Prof. Cas Cremers, Prof. Michael Backes, Dr. Nils Ole Tippenhauer

University of Oxford: Dr. Reuben Binns

University of Torino / ISI Foundation: Prof. Ciro Cattuto

Aix Marseille Univ, Université de Toulon, CNRS, CPT: Dr. Alain Barrat

IMDEA Software Institute: Prof. Dario Fiore

INESC TEC: Prof. Manuel Barbosa (FCUP), Prof. Rui Oliveira (UMinho), Prof. José Pereira (UMinho)

A collaborative (continued) ~~sprint~~ ~~Marathon~~ Ironman

Maintenance and support

March 2020 - **Start**

April 2020 – **GAEN is announced**

May 2020 – **Final version DP3T**

June 2020 – **Pilot SwissCovid (& other EU apps)**

July 2020 – **SwissCovid launch**

August/September 2020 – **Towards international interoperability**

September/November 2020 – **Presence tracing**

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

CISPA: Prof. Cas Cremers, Prof. Michael Backes, Dr. Nils Ole Tippenhauer

University of Oxford: Dr. Reuben Binns

University of Torino / ISI Foundation: Prof. Ciro Cattuto

Aix Marseille Univ, Université de Toulon, CNRS, CPT: Dr. Alain Barrat

IMDEA Software Institute: Prof. Dario Fiore

INESC TEC: Prof. Manuel Barbosa (FCUP), Prof. Rui Oliveira (UMinho), Prof. José Pereira (UMinho)

Technology to help with pandemic contention

- **Manual tracing overwhelmed**

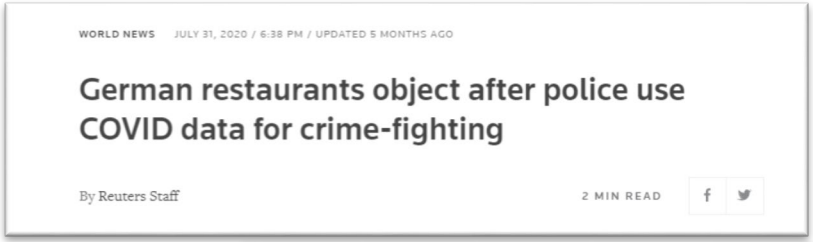
- **The need**
  - **A complement** to **notify** users that have been exposed to COVID19 and they are at risk of infection
  - In a **timely, efficient, and scalable** manner

The constraints: Security and Privacy

- Protect from misuse (surveillance, manipulation, etc)
  - **Purpose limitation by default**

Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design.Computers, Privacy & Data Protection. 2011

The constraints: Security and Privacy



WORLD NEWS JULY 31, 2020 / 6:38 PM / UPDATED 5 MONTHS AGO

German restaurants object after police use COVID data for crime-fighting

By Reuters Staff 2 MIN READ

COVID contact tracing sheet leaves 'creepy' barman to text model

Digital Staff • 7NEWS Published: Saturday, 12 September 2020 3:03 AM

Australia's spy agencies caught collecting COVID-19 app data

Zack Whittaker @zackwhittaker / 4:32 PM GMT+1 • November 24, 2020 Comment

Covid 19 coronavirus: Subway worker 'harassed' woman customer after getting details for contact tracing

14 May, 2020 08:23 PM 3 minutes to read

BBC Sign in Home News Sport Reel Worklife Travel Future Cultu

NEWS January 2021

Home | Coronavirus | Video | World | UK | Business | Tech | Science | Stories | Entertainment & Arts | Health

Asia | China | India

Singapore reveals Covid privacy data available to police

Top Stor

Indonesi

The constraints: Security and Privacy

- Protect health-related data
- Protect from misuse (surveillance, manipulation, etc)
  - **Purpose limitation by default**
  - hide users identity, location, and behavior (social graph)

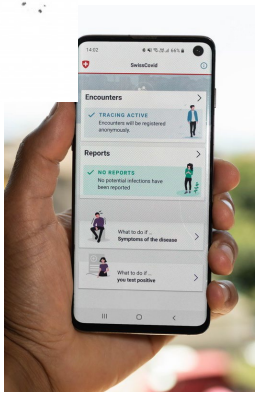
Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design.Computers, Privacy & Data Protection. 2011

The constraints: Security and Privacy

- Protect health-related data
- Protect from misuse (surveillance, manipulation, etc)
  - **Purpose limitation by default**
  - hide users identity, location, and behavior (social graph)
- Preserve system integrity
  - Prevent false alarms & Denial of Service

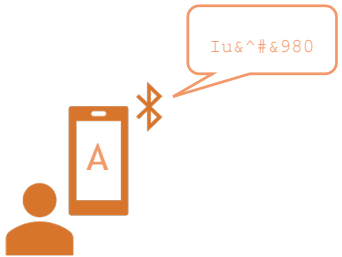
Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design.Computers, Privacy & Data Protection. 2011

The “hidden” constraint Reality



- High scalability and reliability
- Design under time pressure!
  - Need fast, robust verification
    - KISS principle: Keep It Simple Stupid
    - Avoid new technologies or non-mainstream
  - Use existing infrastructure
    - BLE beacons
- Dependencies, dependencies, dependencies

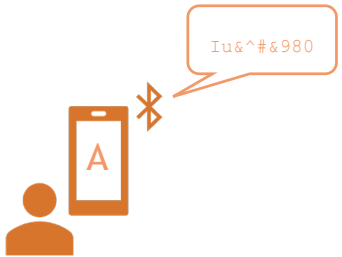
The system design
Our first idea



- The App creates a **secret key (SK**) and from this key it derives **random identifiers (EphIDs)** that it broadcasts via Bluetooth
- Secret keys are rotated every day $SK_{t+1}=H(SK_t)$
- $EphID_1 \,||\, ... \,||\, EphID_n = PRG(PRF(SK_t,\text{“broadcast key”}))$
- A random identifier is used for a limited amount of time
- Without the key, no-one can link two identifiers

https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

The system design
First quicksand pond...



- The App creates a **secret key (SK)** and from this key it derives **random identifiers (EphIDs)** that it broadcasts via Blueto
- Secre
  SK_{t+1}=
- $EphID_1$ | st key"))
- A ranc ted amour
- Without the key, no-one can link two identifiers

https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

Reality
Use existing infrastructure

- Battery and CPU usage
  - Limited round trips
  - Google and Apple **must** be involved
- Run in the background
  - Apple **must** be involved
- Compatibility Android - iOS
  - Google and Apple **must** be involved

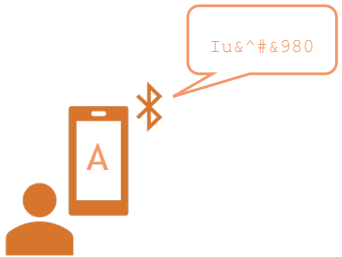
- Google and Apple implement the protocol **and the API**
  - Implications on privacy engineering
  - Implications for epidemiology and exposure estimation (no time in this talk…)
  - Implications for privacy when internationalizing (no time in this talk…)

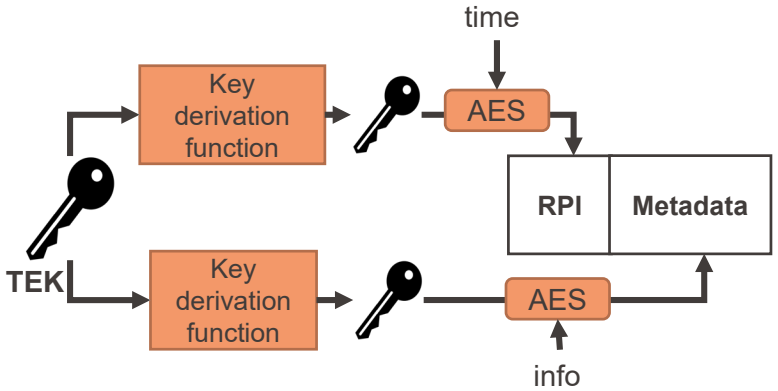
Blagovesta Kostova, Seda Gürses, Carmela Troncoso. Privacy Engineering Meets Software Engineering. On the Challenges of Engineering Privacy By Design

Carmela Troncoso

The system design Platform decides Exposure Notification

- The App creates a **secret every day** (**TEK**) and from this key it derives **random identifiers (RPIs)** that it broadcasts via Bluetooth

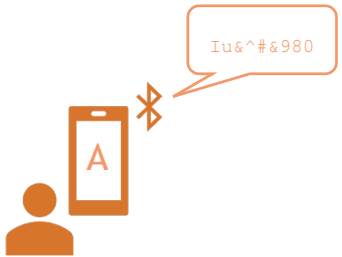




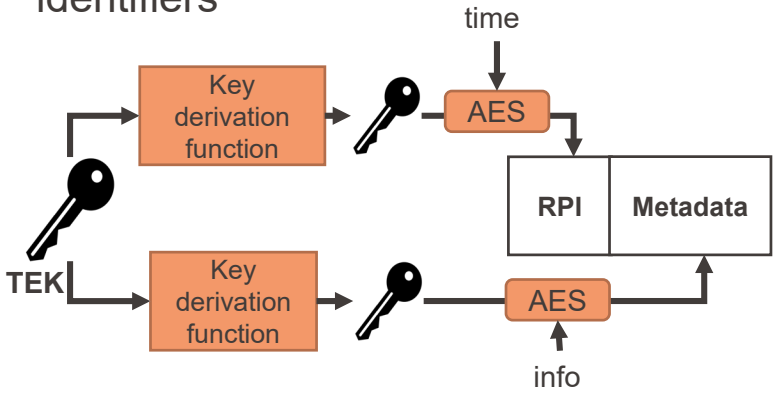
https://blog.google/documents/69/Exposure_Notification_-_Cryptography_Specification_v1.2.1.pdf

Carmela Troncoso

The system design Platform decides Exposure Notification

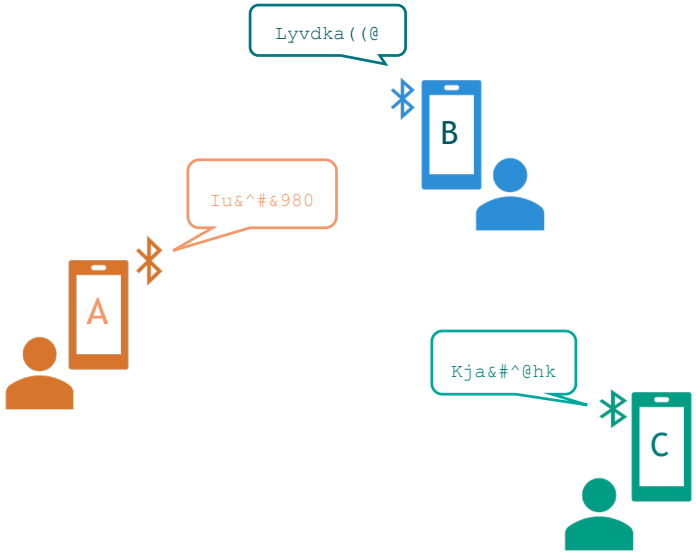


- The App creates a **secret every day** (**TEK**) and from this key it derives **random identifiers (RPIs)** that it broadcasts via Bluetooth
- A random identifier is used for a limited amount of time
- Without the key, no-one can link two identifiers

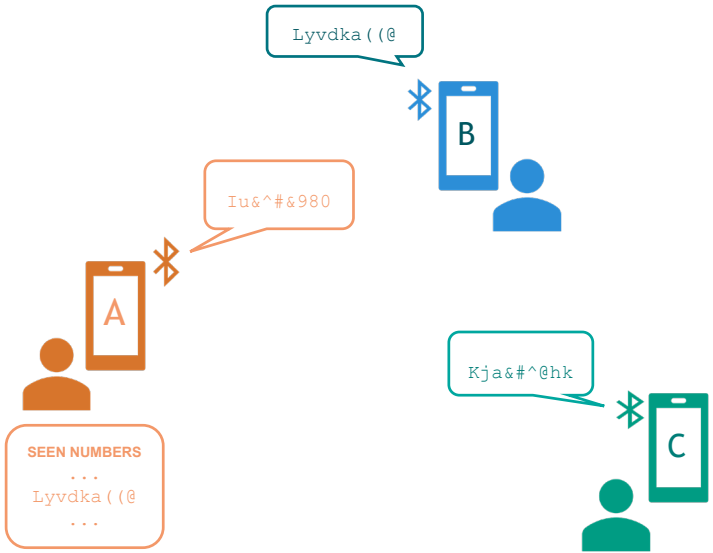


https://blog.google/documents/69/Exposure_Notification_-_Cryptography_Specification_v1.2.1.pdf

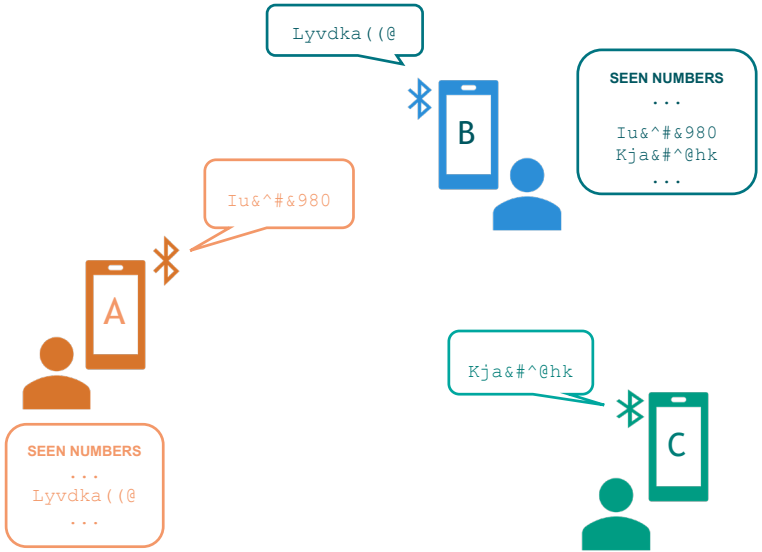
The system design



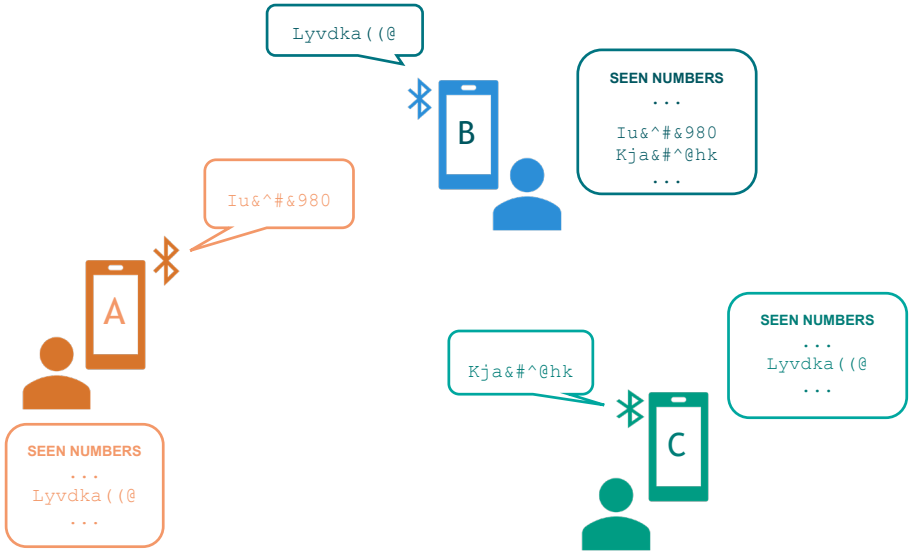
The system design



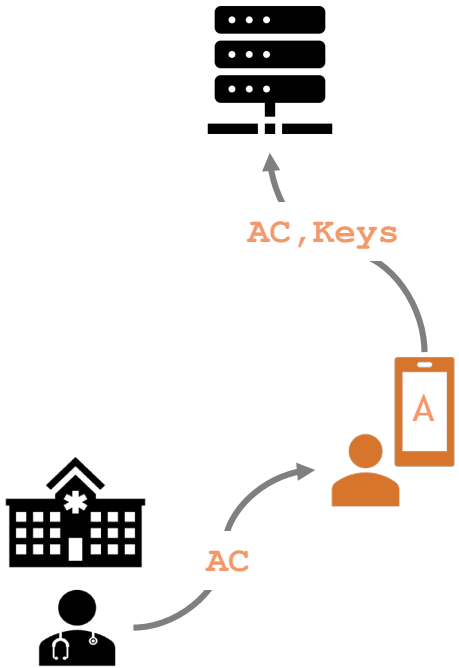
The system design



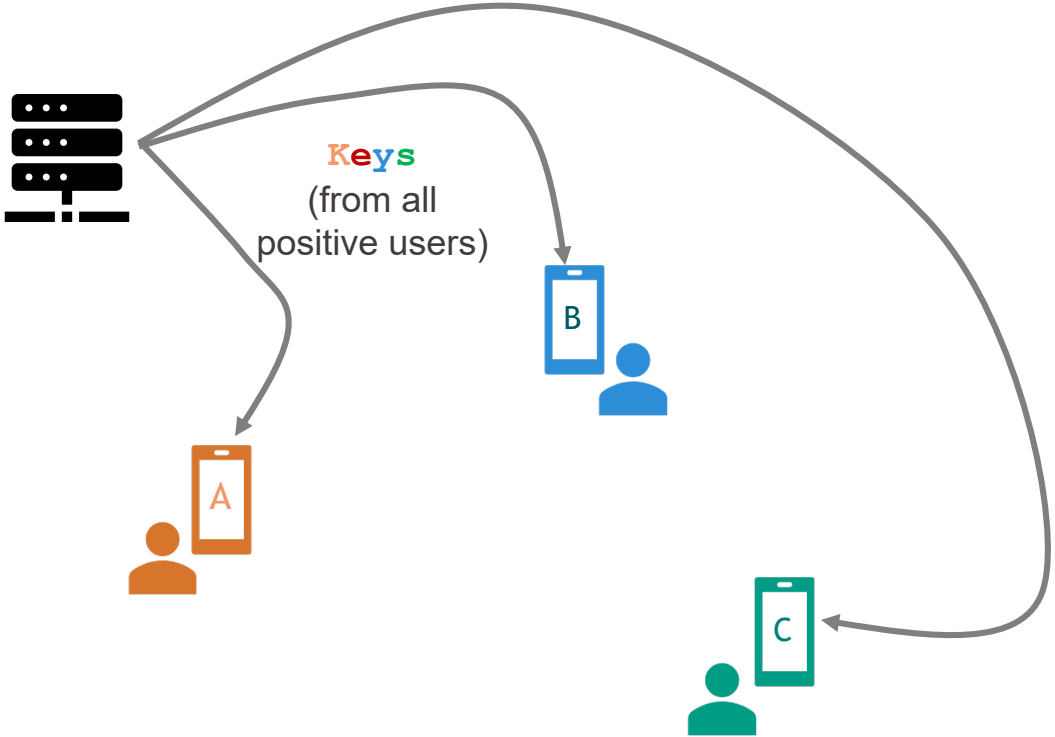
The system design



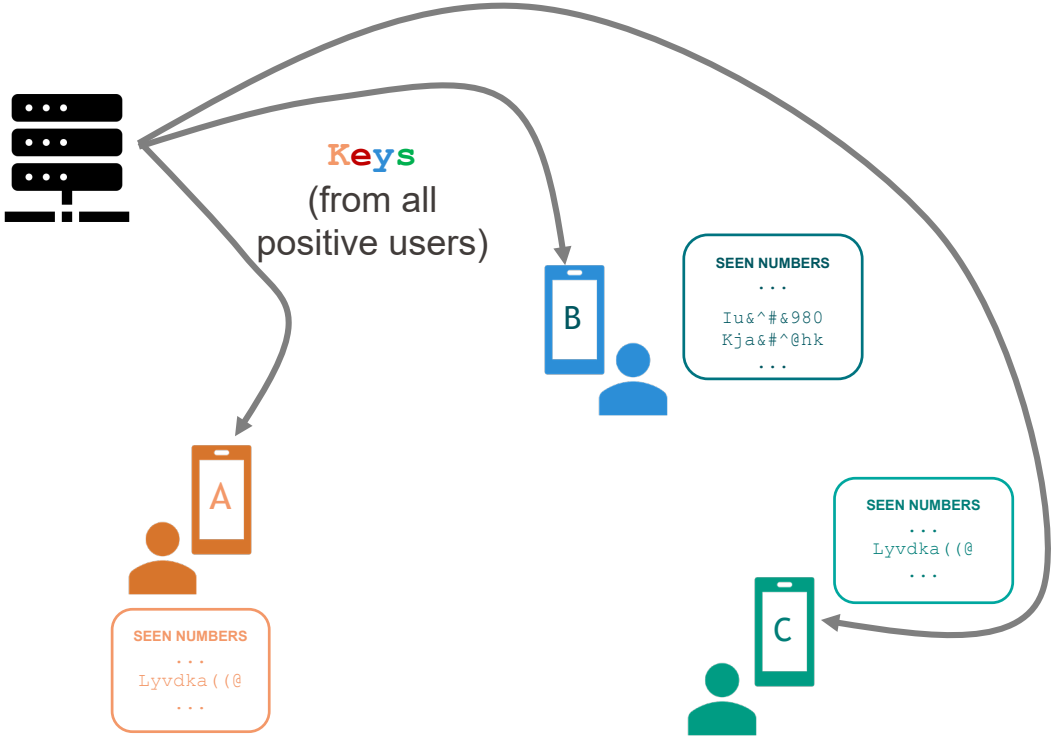
The system design



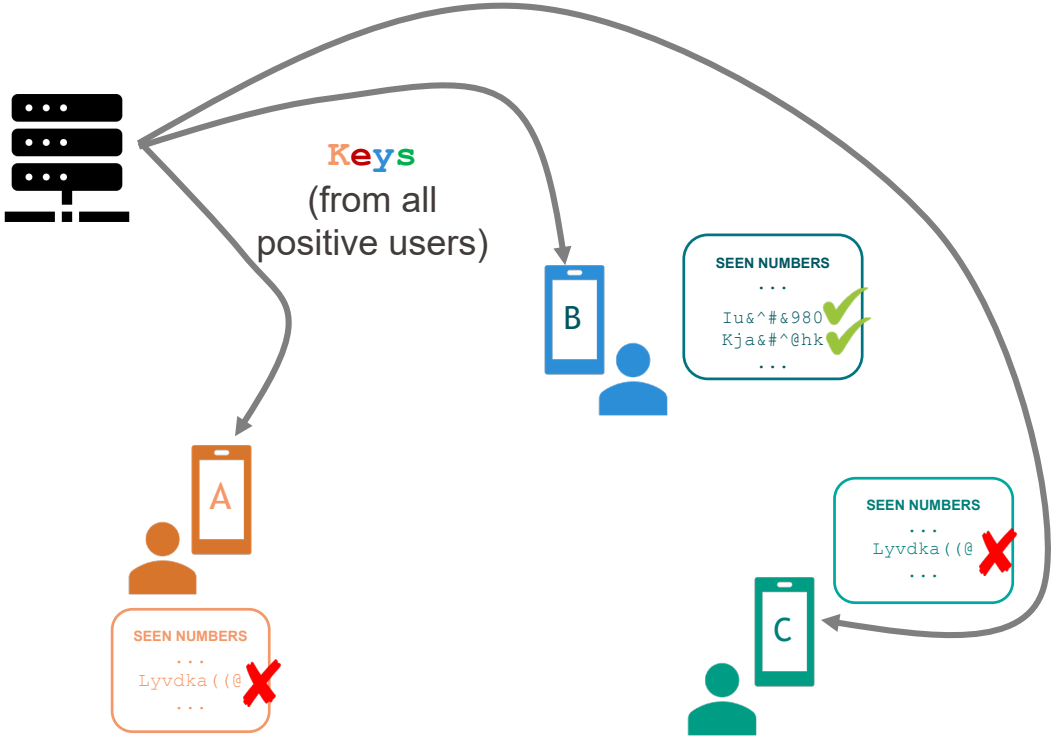
The system design



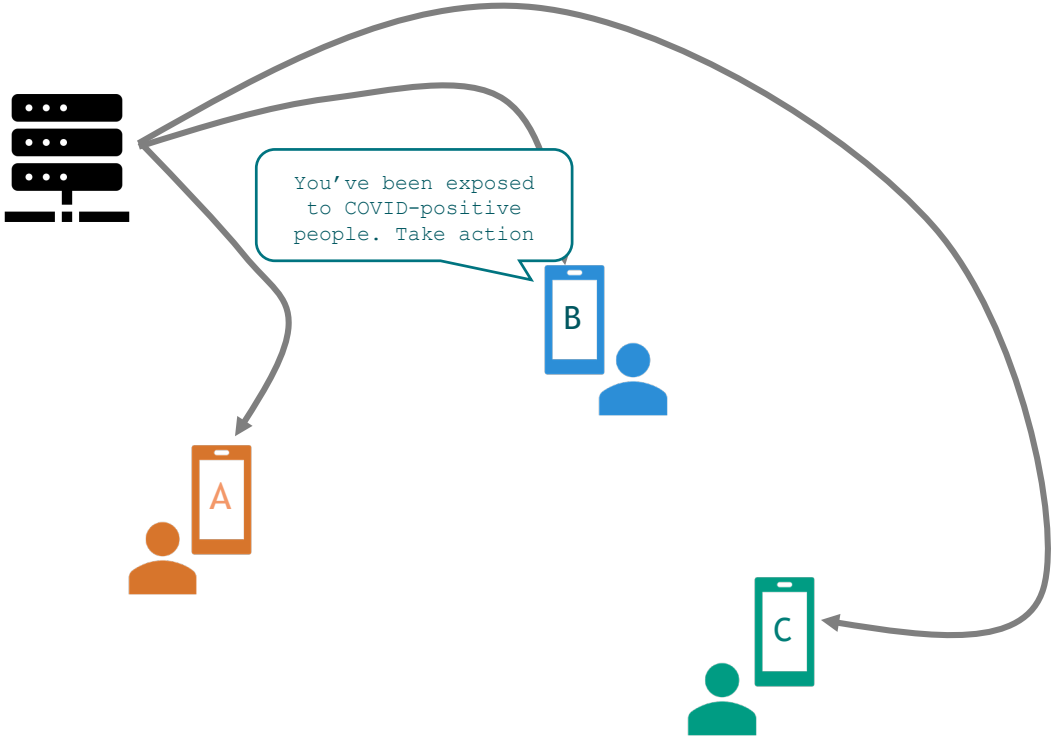
The system design



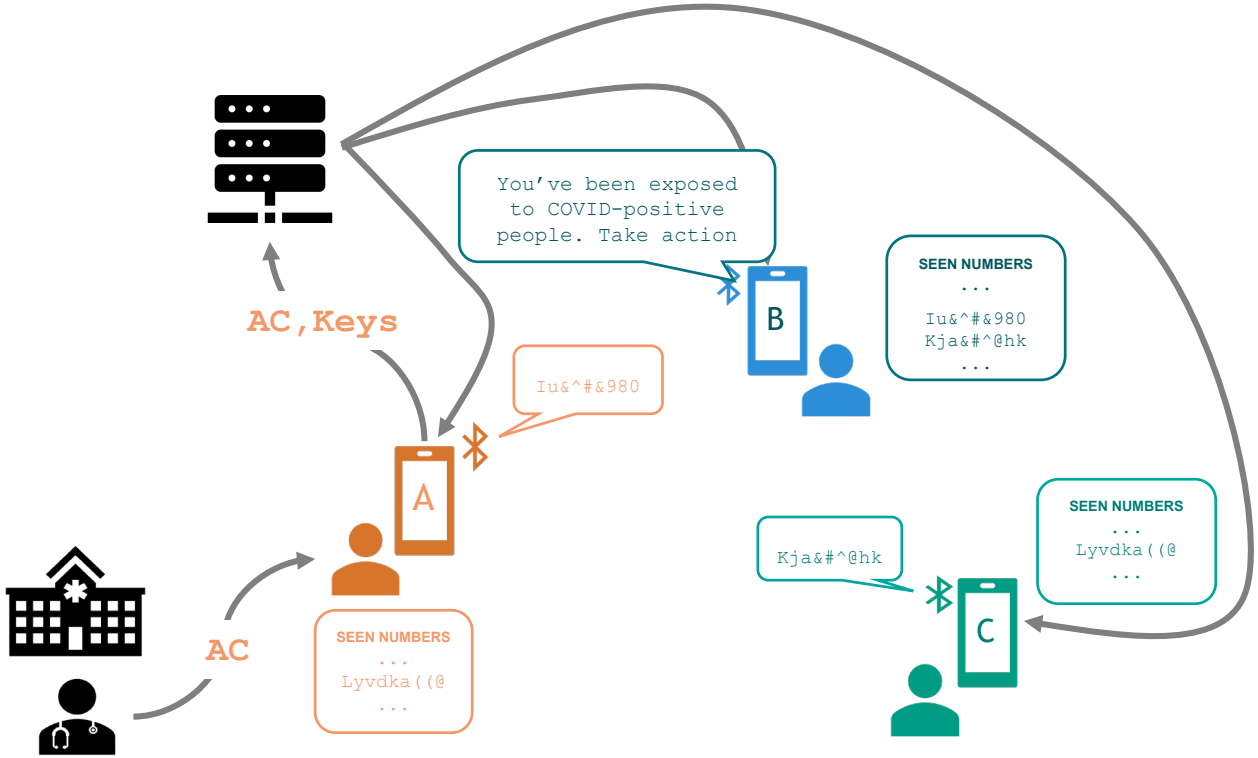
The system design



The system design



The system design



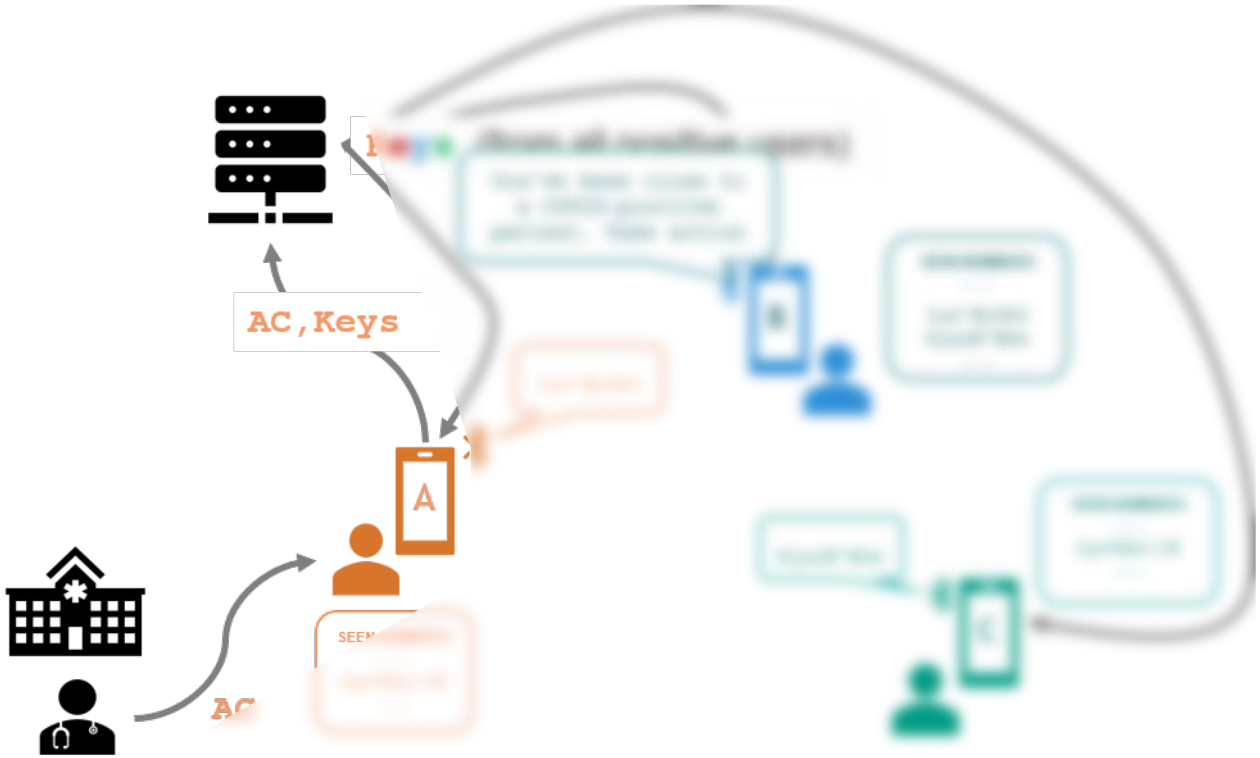
Only information that ever leaves the phone are the **TEKs broadcasted** during the contagious period.

No identity, **no** location, **no** information about others

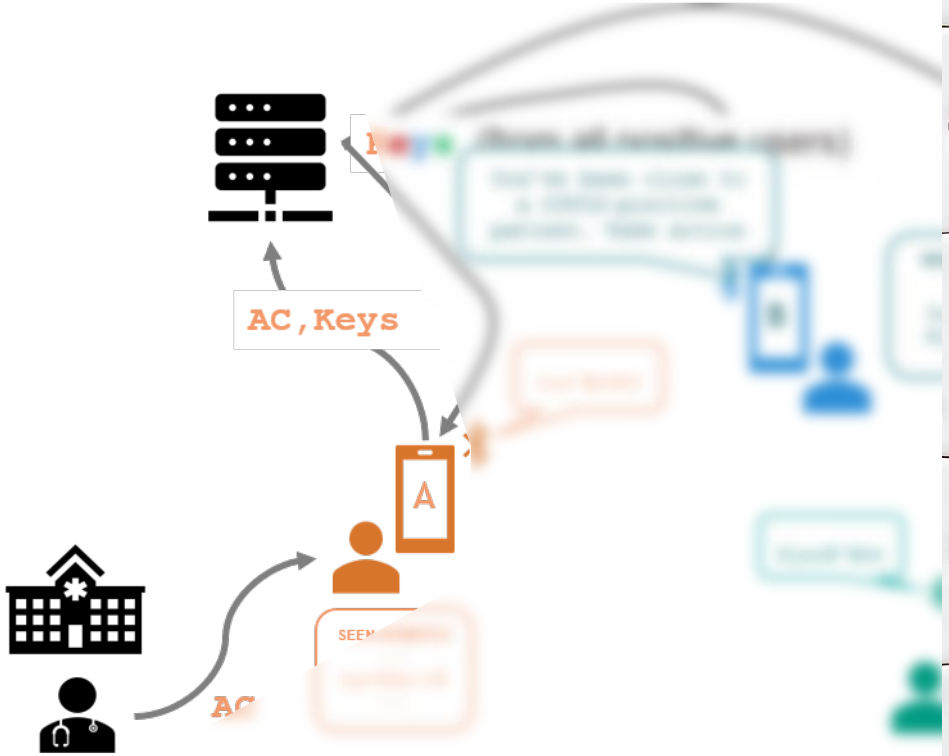
No information available for abuse

System **sunsets-by-design**

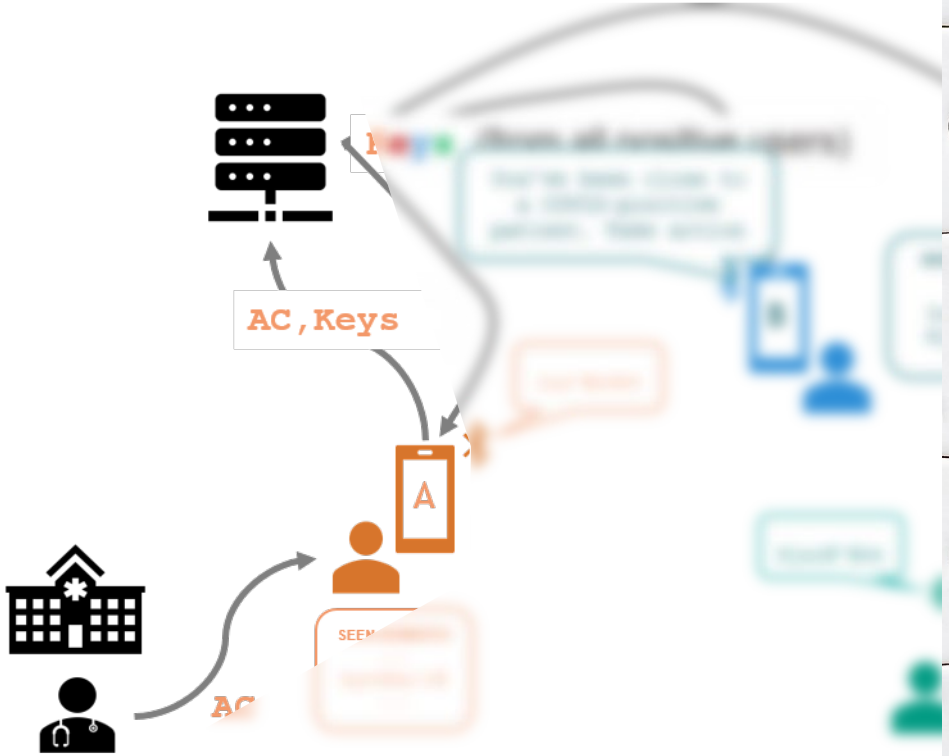
The system design



The system design

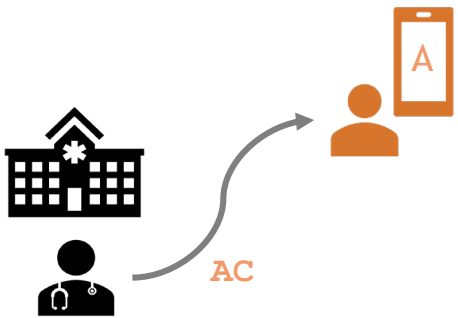


The system design



Authorization mechanism
Our first design

- Crucial for security: only *true* positives can upload
  - Desired properties:
    - Privacy
    - Hard to delegate
  - Crypto FTW! Commit to content in authorization token!

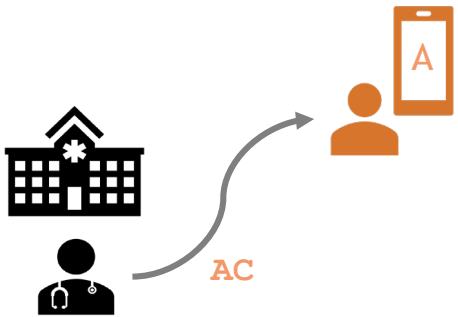


Secure Upload Authorisation for Digital proximity tracing. DP-3T project. https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Upload%20Authorisation%20Analysis%20and%20Guidelines.pdf

Carmela Troncoso

Authorization mechanism

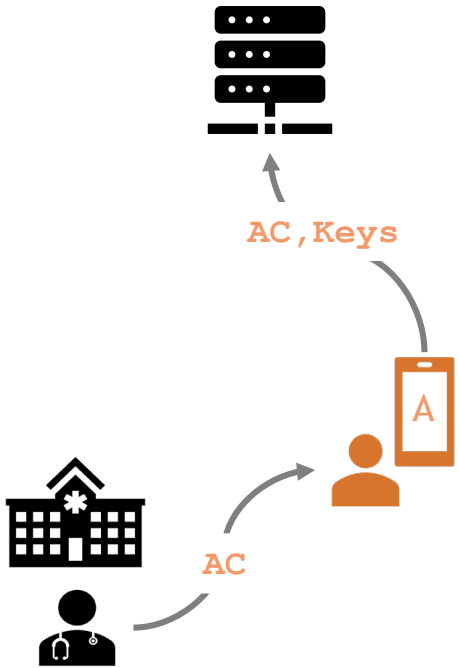
- Crucial for security: only *true* positives can upload
  - Desired properties:
    - Privacy
    - Hard to delegate
  - Crypto FTW! Commit to content in authorization token!



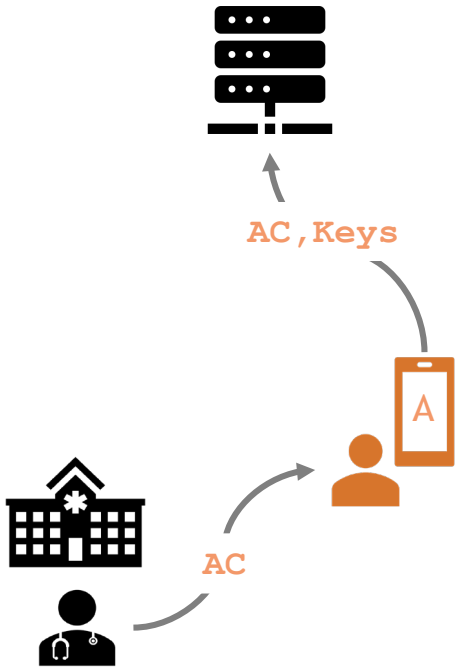
- Health systems/staff are not digitalized everywhere
  - Simple activation codes sent via phone/mail/sms
  - Different level of automatization
  - Belgium went for (light) commitments!

Secure Upload Authorisation for Digital proximity tracing. DP-3T project. https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Upload%20Authorisation%20Analysis%20and%20Guidelines.pdf

Privacy engineering
Are we done?



Privacy engineering
Are we done?

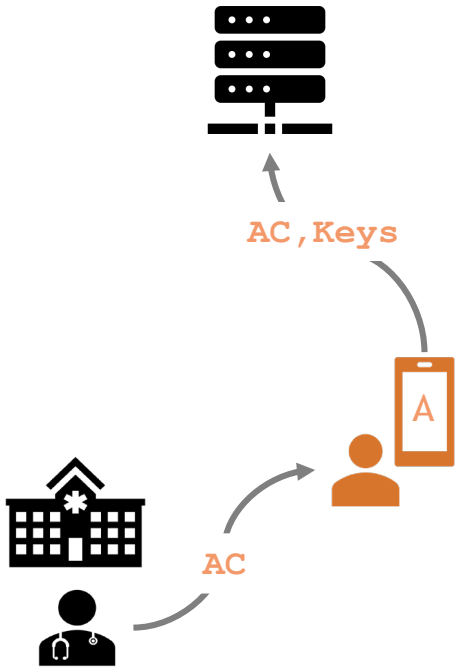


Existence of upload

⇩

the user is COVID+

Privacy of uploads
Our first idea



Existence of upload

⇩

the user is COVID+

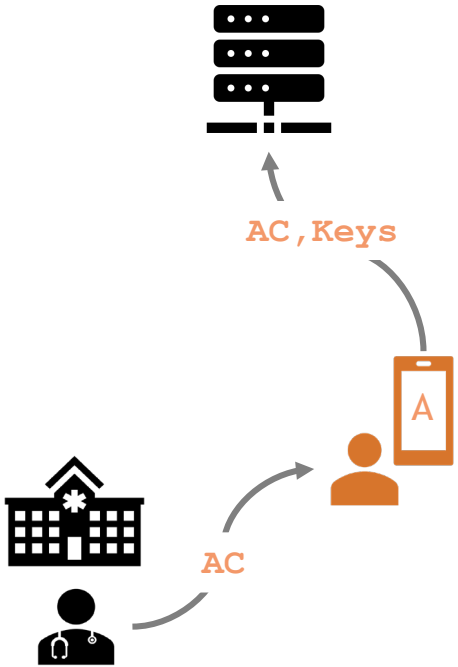
DP3T design paper

The pattern associated with the upload of identifiers to the server would reveal the COVID-19 positive status of users to network eavesdroppers (ISP or curious WiFi provider) and tech-savvy adversaries. If these adversaries can bind the observed IP address to a more stable identifier such as an ISP subscription number, then they can de-anonymize the confirmed positive cases. This can be mitigated by using dummy uploads. These

- https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

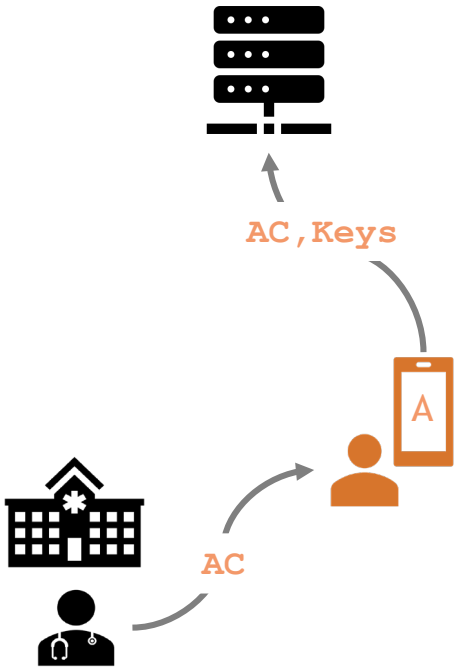
Privacy of uploads Practice

- Unknown environment
  - What is users' behavior?



- https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Privacy of uploads Practice

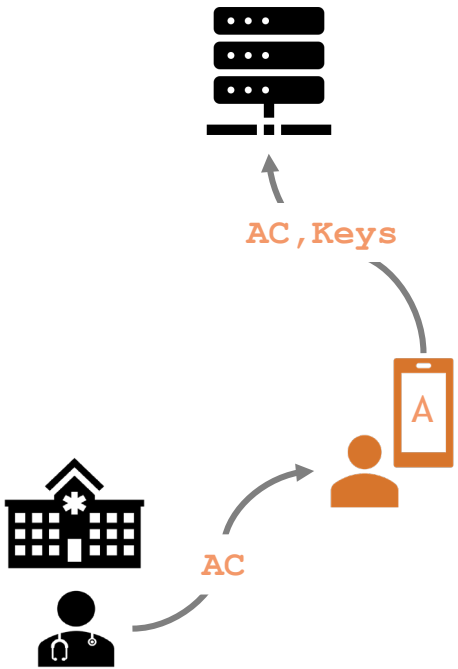


- Unknown environment
  - What is users' behavior?

- Constraints associated to the platform
  - Bandwidth
  - Server capacity
  - Battery

- https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Privacy of uploads Practice



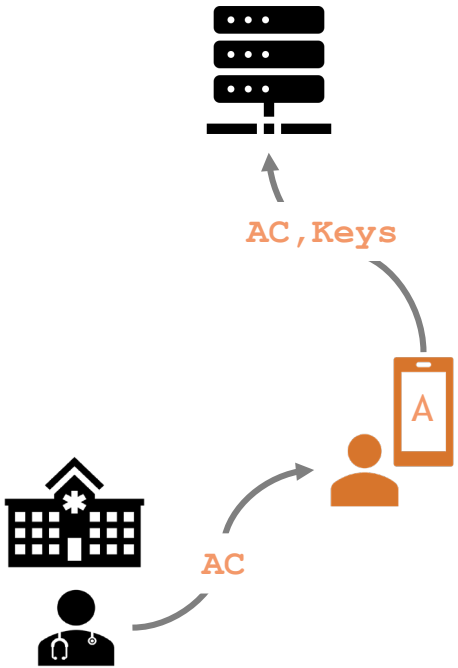
- Unknown environment
  - What is users' behavior?

- Constraints associated to the platform
  - Bandwidth
  - Server capacity
  - Battery

- Anonymity and delays not possible

- https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Privacy of uploads Practice



- Unknown environment
  - What is users' behavior?

- Constraints associated to the platform
  - Bandwidth
  - Server capacity
  - Battery

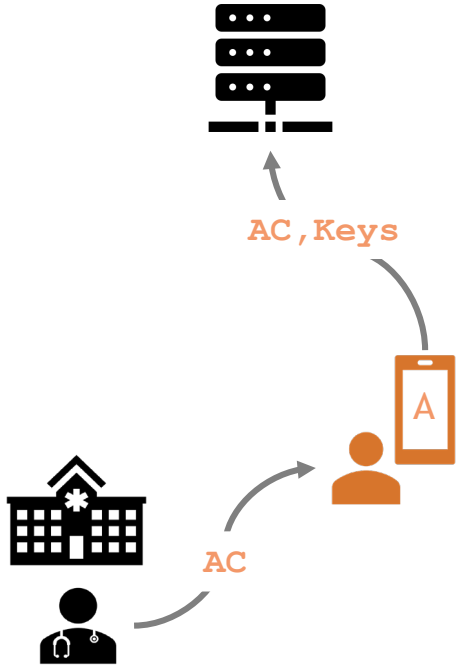
- Anonymity and delays not possible

⇩

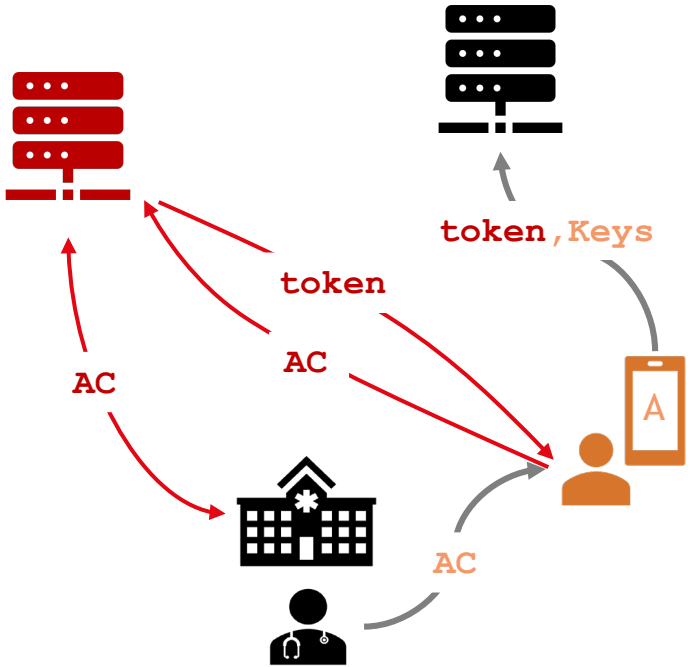
Plausible deniability
(constant time & size)

- https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Privacy of uploads
Practice – there is authentication!

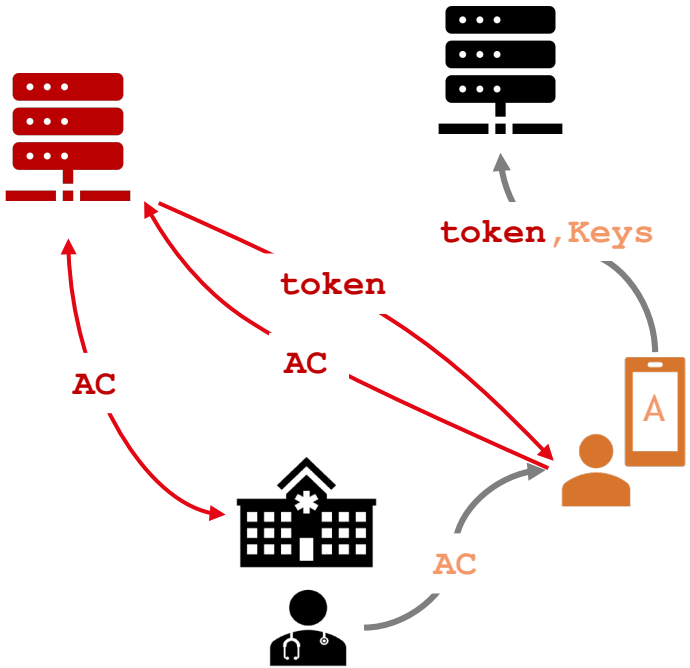


Privacy of uploads
Practice – there is authentication!



- https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

Privacy of uploads
Practice – there is authentication!

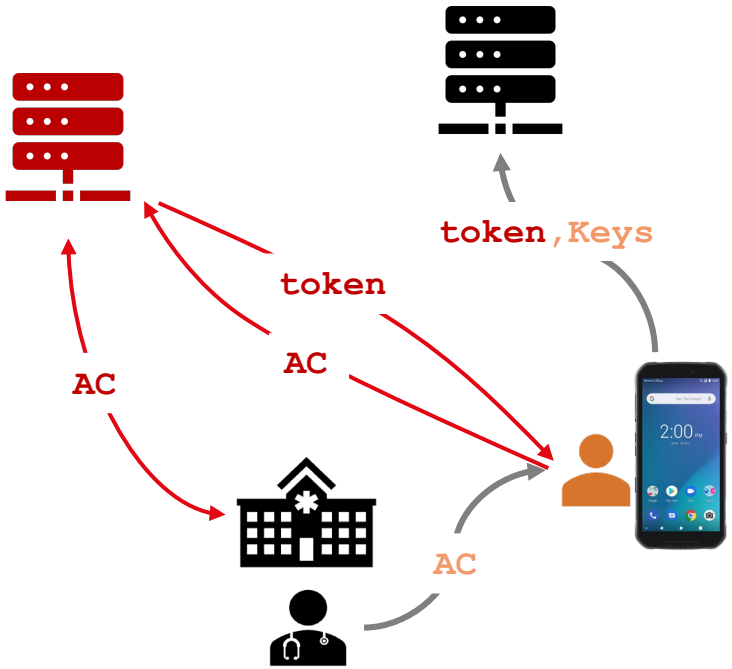


- Dummies also must realize the authentication step
  - Servers must consider dummies
  - Ensure equal timing and volume

- https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

Privacy of uploads Practice –

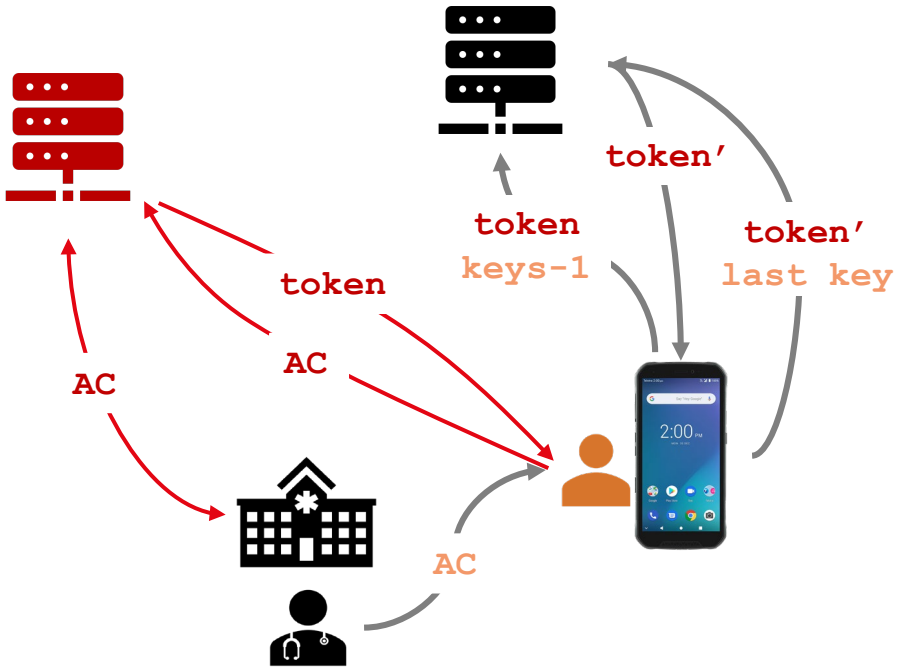
- Exposure Notification API (<v1.5) had one security mechanism:
  - Only reveal key after it expires
  - (Not needed, it is an implementation decision)



- Implications on authorization and dummy strategy
  - Cannot delay all keys!

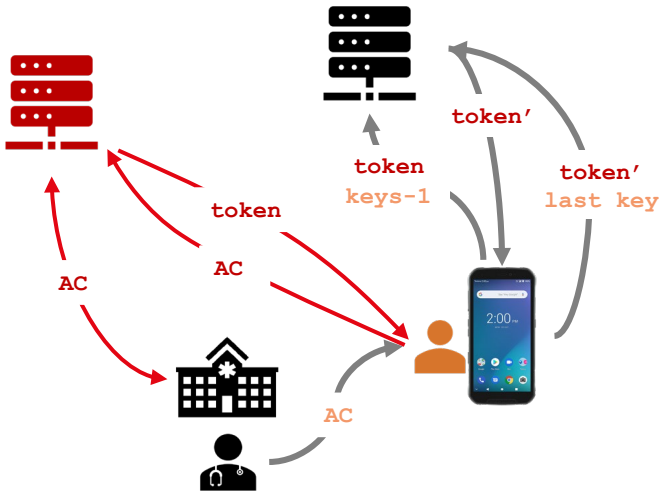
Privacy of uploads Practice –

- Exposure Notification API (<v1.5) had one security mechanism:
  - Only reveal key after it expires
  - (Not needed, it is an implementation decision)

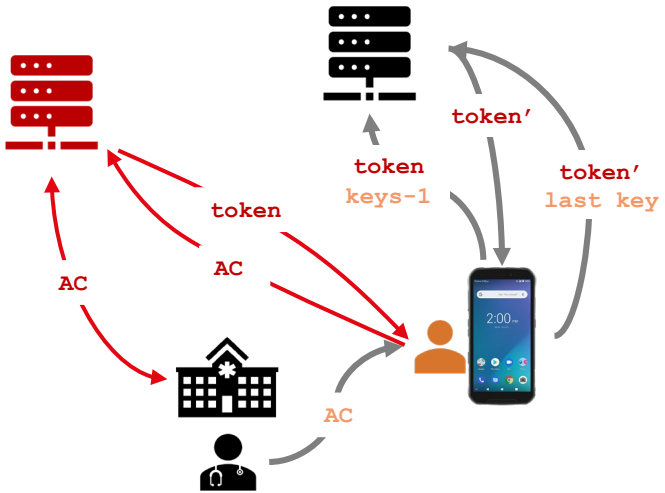


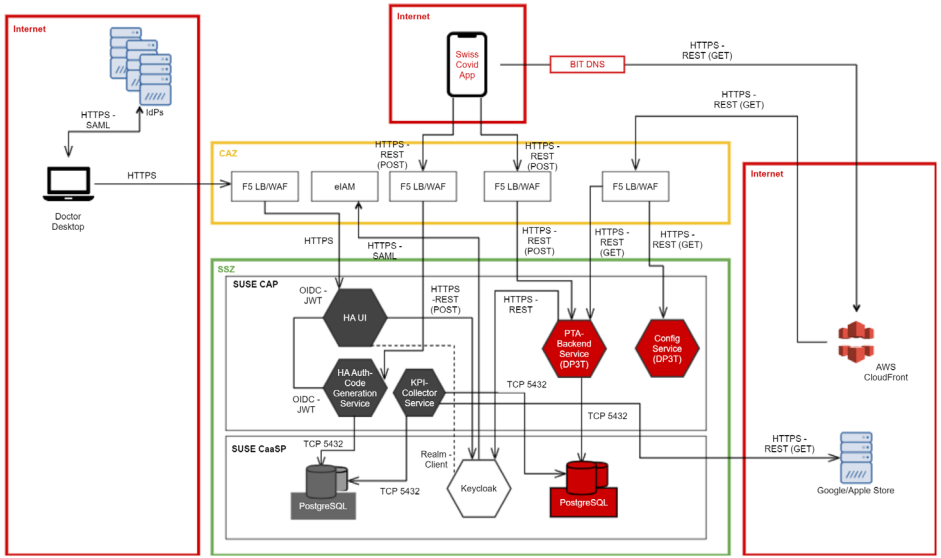
- Implications on authorization and dummy strategy
  - Cannot delay all keys!
  - Dummies must mimic second upload

Privacy of uploads
Practice – servers don't exist in the vacuum



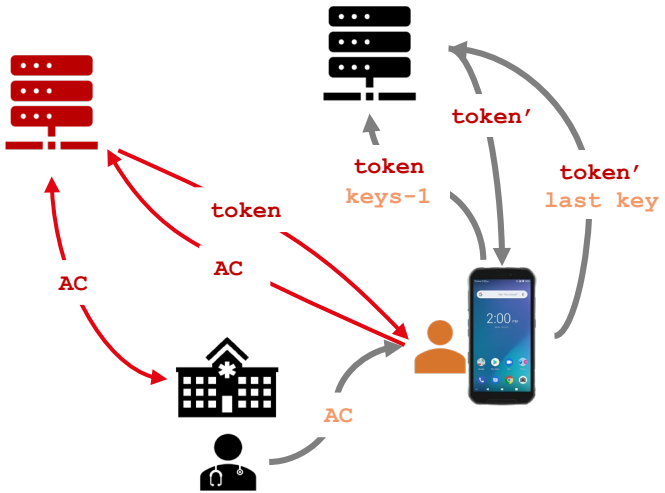
Privacy of uploads
Practice – servers don't exist in the vacuum

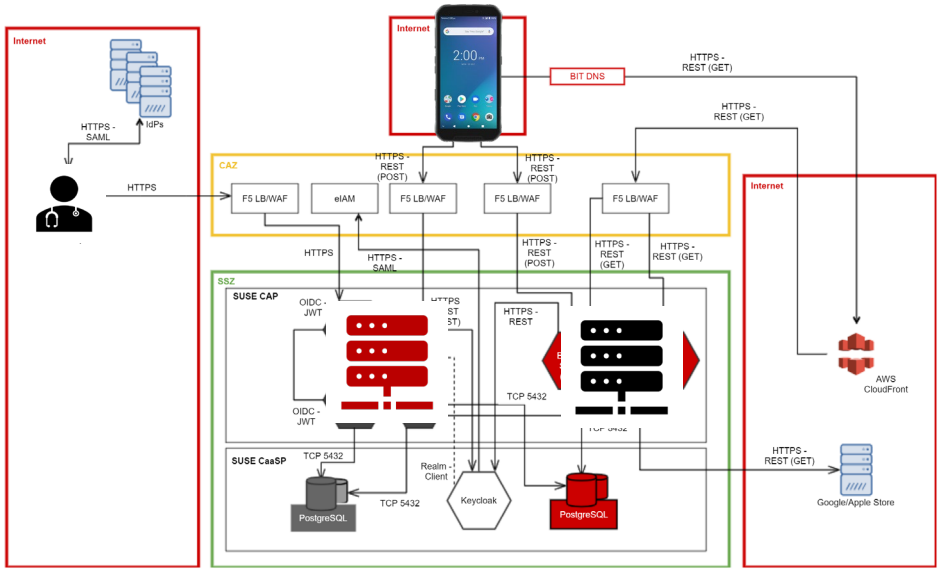




- https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Privacy of uploads
Practice – servers don't exist in the vacuum

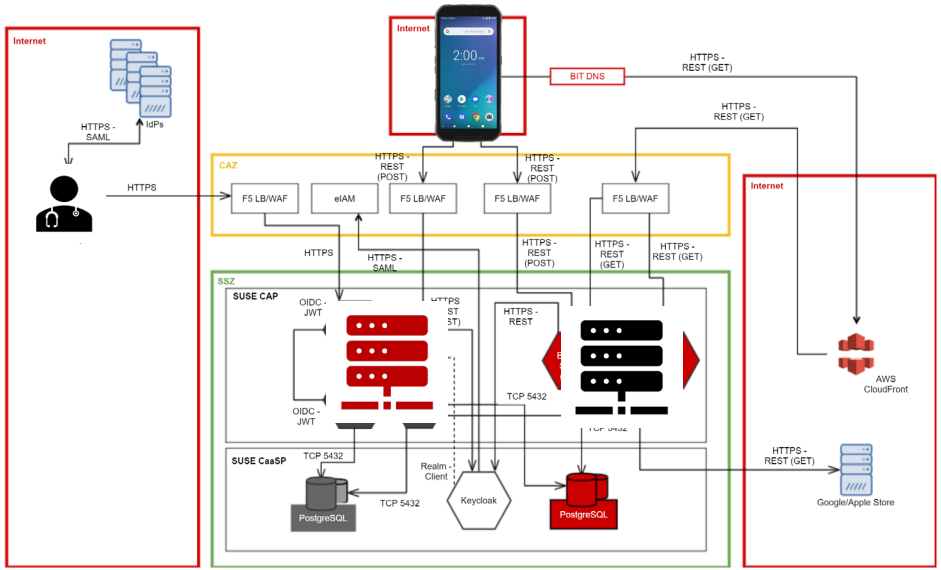




- https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Privacy of uploads
Practice – servers don't exist in the vacuum

- Load Balancer, Firewall
  - More information than expected!
  - Off the shelf cloud managing tools
- Careful design of logging to avoid forensics
  - Coarse logging at key server
  - Only counts logged for statistics
    - e.g, active users based on dummy traffic
- Logging strategy re-designed N times



https://github.com/DP-3T/documents/blob/master/DP3T%20-%20Best%20Practices%20for%20Operation%20Security%20in%20Proximity%20Tracing.pdf

Where is this deployed?

1.87 Million active users (~22% population)

~18000 COVID-positive users uploaded their keys in December (15% of PCR in Switzerland)

Field experiment in Zurich October 2020

- 80% COVID-positive app users upload their codes
- 22% sent quarantine
- 1 in 10 tested positive after notification
- 5% of positives with respect to Manual Contact Tracing in Zurich
- Speed: ~1 day faster notification for non-household exosures (70% of the cases)

https://www.experimental.bfs.admin.ch/expstat/en/home/innovative-methods/swisscovid-app-monitoring.html
https://github.com/digitalepidemiologylab/swisscovid_efficacy/blob/master/SwissCovid_efficacy_MS.pdf
https://www.ebpi.uzh.ch/dam/jcr:5fc56fb7-3e7e-40bf-8df4-1852a067a625/Estimation%20of%20SwissCovid%20effectiveness%20for%20the%20Canton%20of%20Zurich%20in%20September%202020_V1.5.pdf
https://www.medrxiv.org/content/10.1101/2020.12.21.20248619v1.full.pdf

Key lessons

- Data is not a must!
- Privacy engineering goes well beyond crypto
- Privacy engineering in an agile/service world is exhausting
  - Platforms and requirements continuously change
- Good socio-technical integration is key to success and it is **hard**
  - Purpose limitation and abuse prevention is a must