Infosec aspects of smartphone contact tracing applications

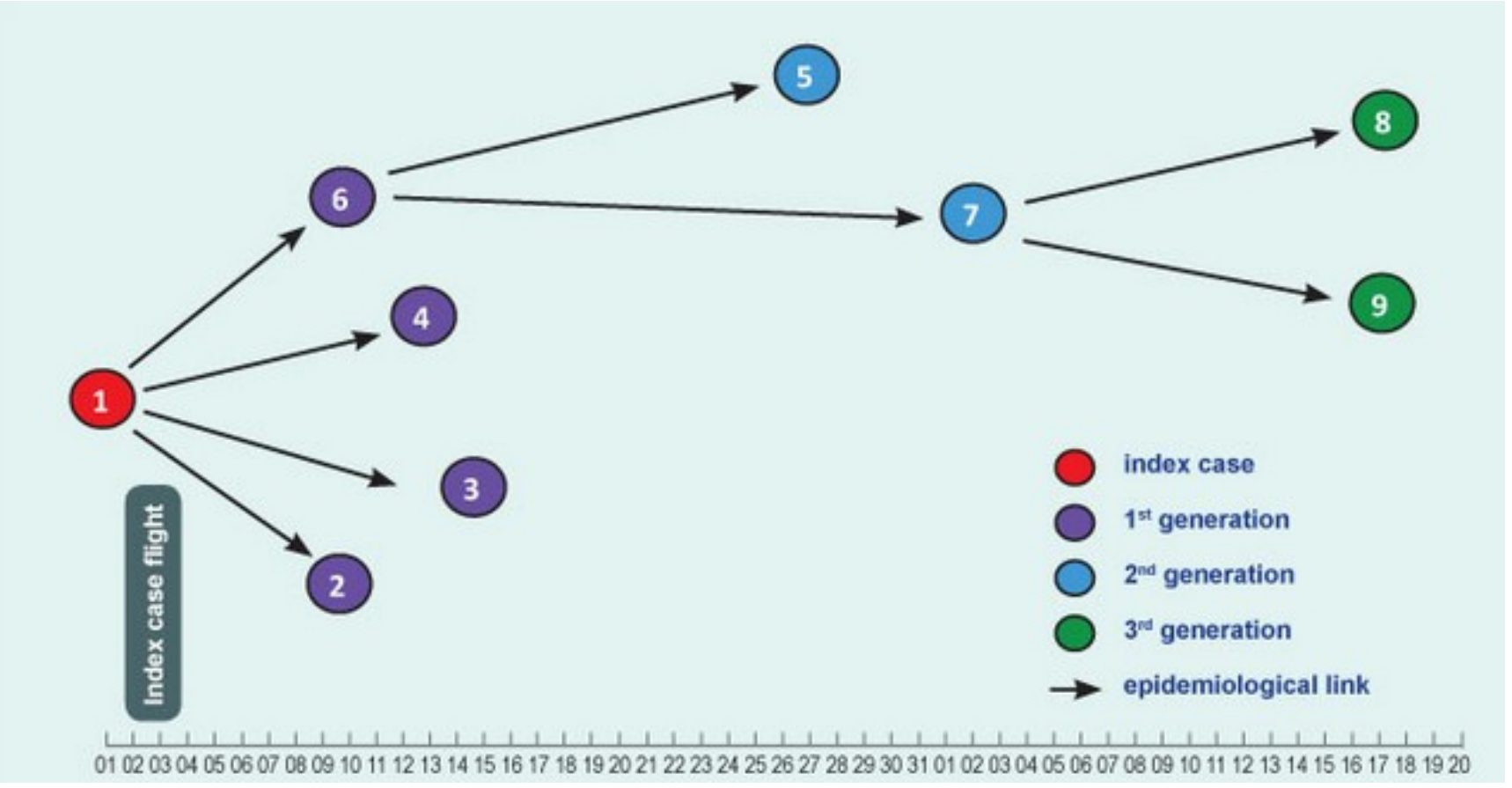
George Magklaras
Lucia Lopez Bojorquez

International Symposium on Human Aspects of Information Security & Assurance
8-10 July 2020
HAISA 2020

Agenda

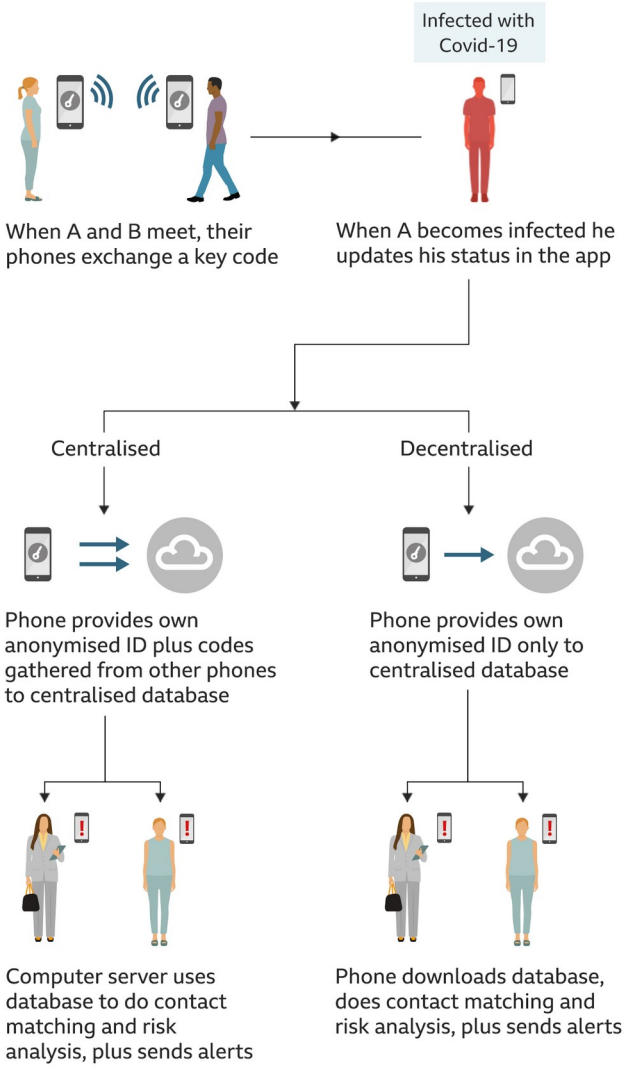
- Contact tracing
- Digital contact tracing
- Smartphone contact tracing issues
  - GPS/Bluetooth
  - Bluetooth issues
- Central IT infrastructure/data management issues

Contact tracing



Digital contact tracing

- Manual/conventional procedures are:
  - Error prone/slow
  - **Resource intensive**
  - **Non ubiquitous**
- Digital procedures are:
  - Error prone/fast
  - **Lower resourcing requirements**
  - **Ubiquitous**
- Before, during and past COVID-19 times



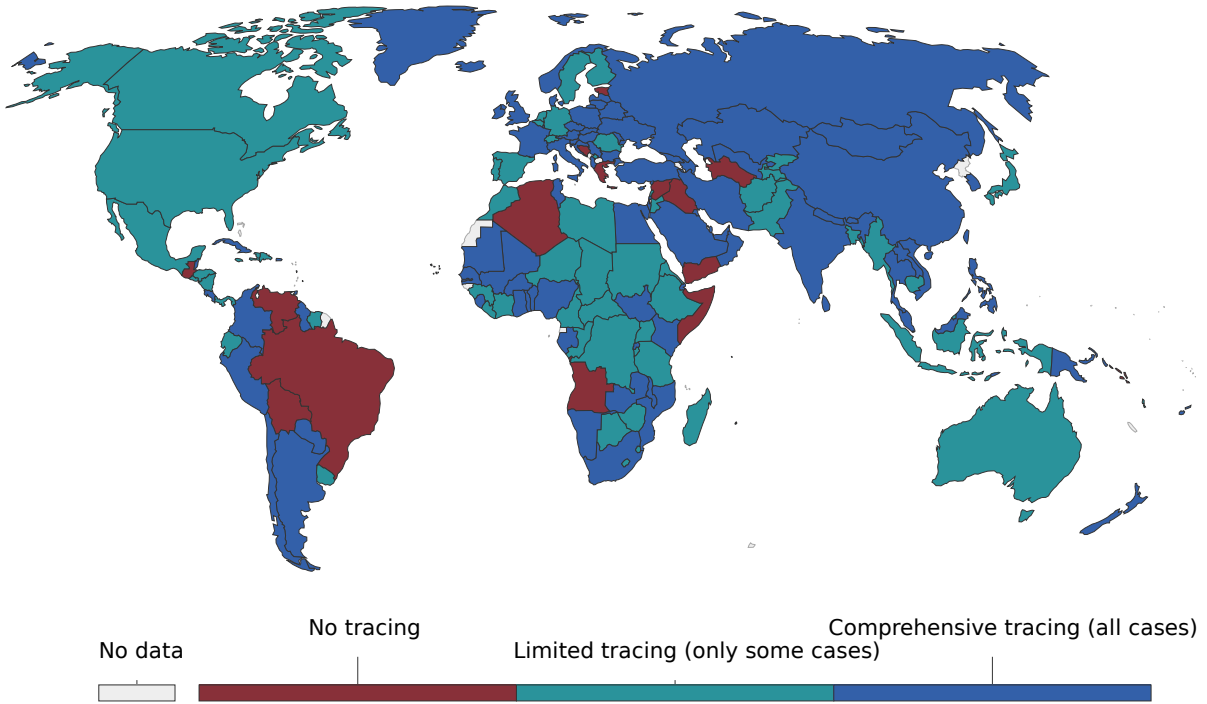
- Proximity sensing component
- Ephemeral anonymous identity
- Integration/inclusion with/to government/health authorities
- Centralized versus decentralized contact tracing processing

Lukewarm response

Which countries do COVID-19 contact tracing?, Jul 5, 2020

Our World in Data

'Limited' contact tracing means some, but not all, cases are traced. 'Comprehensive' tracing means all cases are traced.



Source: Hale, Webster, Petherick, Phillips, and Kira (2020). Oxford COVID-19 Government Response Tracker – Last Updated 6th July.

Major early implementations in:

- Australia (Covidsafe)
- China
- India (Aarogua Setu)
- Israel (The Shield, SAFE)
- Norway (Smittestopp)
- Singapore (TraceTogether)
- South Korea

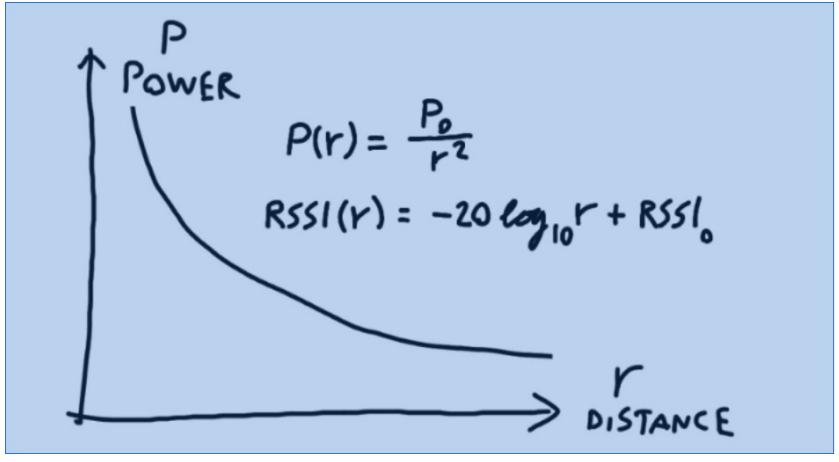
Why the lukewarm response?

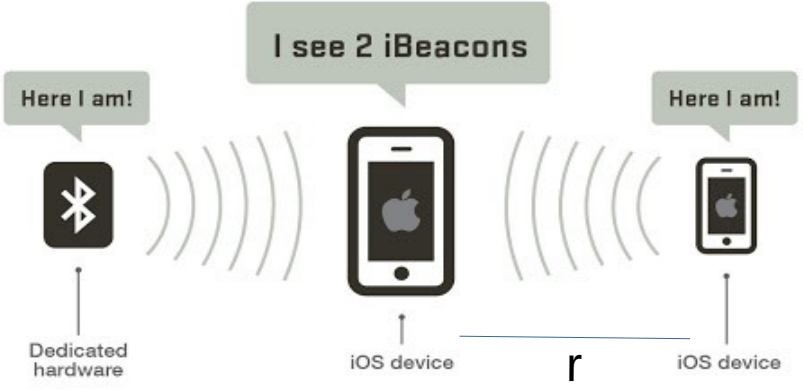
- Rushed implementations as urgent measures
- To what extent these first implementations:
  - Can provide reliable proximity data
  - Can protect the safety of the mobile OS
  - Respect the privacy of the user
  - Manage the collected data in a responsible manner

Proximity sensing vulnerabilities

- Low energy Bluetooth (BLE)
- Does it give reliable proximity sensing data?
- Is user anonymity really preserved?
- Is smartphone security an issue?

BLE proximity sensing





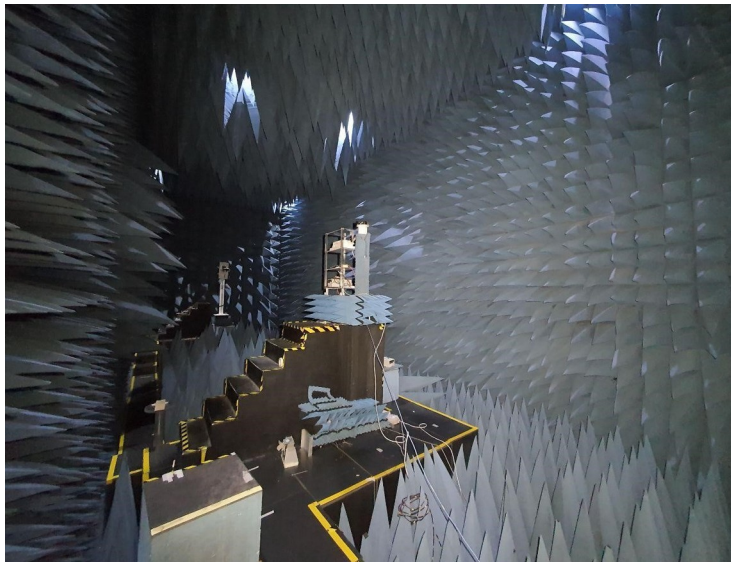
- Inverse square law attenuation
- **Device dependent** (different chipsets,implementations)
- RSSI: a difference of 20 units in dBm means an R estimate multipled by 10
- Sensed at the **receiver**
- Ideal conditions

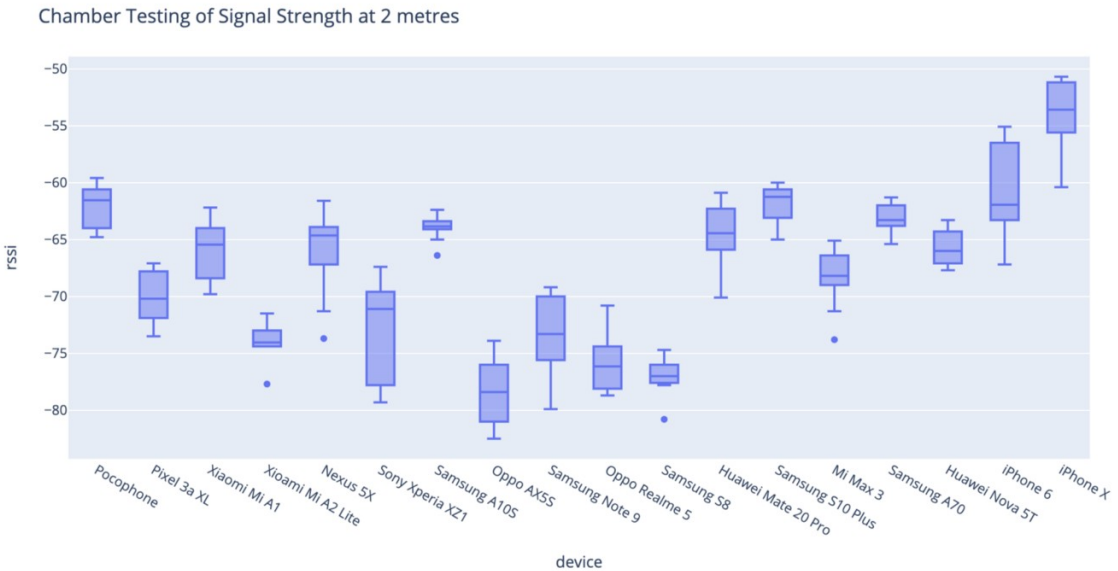
TABLE I. ESTIMATED DISTANCE AND ERROR

Sampling Times	RSSI (dBm)	Estimated Distance (m)	Error (m)
1	-70.2	4.425	1.425
2	-74.2	6.245	3.245
3	-68.1	3.688	0.688
4	-64.8	2.772	0.228
5	-67.6	3.531	0.531
6	-73.1	5.684	2.684
7	-70.4	4.502	1.502
8	-67.9	3.624	0.624
9	-67.3	3.441	0.441
10	-71.1	4.784	1.784
Data from the Nexus 5 as used in Fig. 1		Average Error	1.315 (m)
		Standard Deviation	1.016 (m)

Based on the above results, 3 meters away from the beacon can lead to an average distance estimation error of 1.315 meters, which in our case can mislead the user to a wrong bookshelf. Furthermore, this pre-experiment was just carried in a free space area without considering the signal attenuation caused by human body absorption, the movement of people in real time, and the complicated layout of the space in the library with walls and metal bookshelves. Hence, we abandoned using the Signal Path Loss Model.

BLE proximity sensing (2)

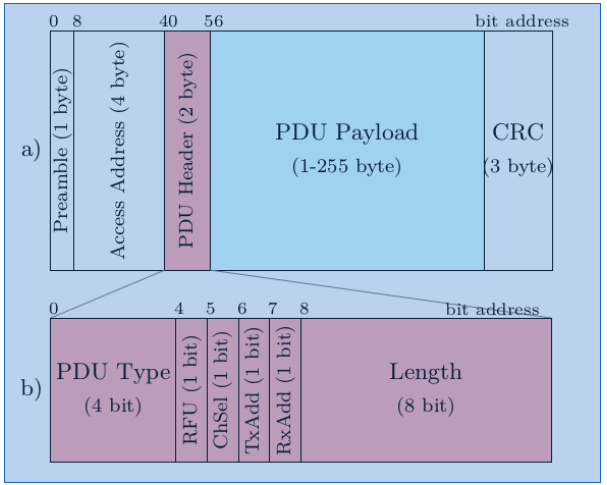


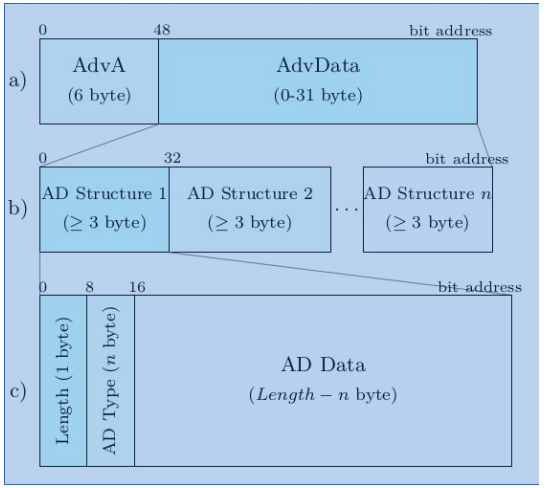


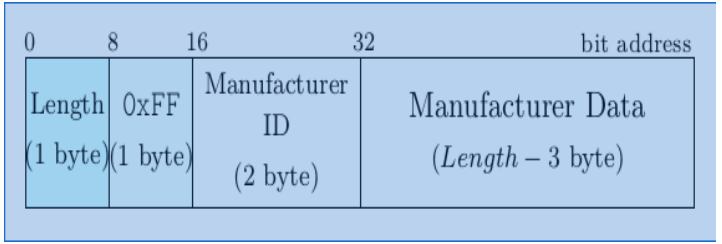
Source: OpenTrace community github repo

- Differences in RSSI require complex smartphone calibration
- In the real world, BLE is a noisy protocol
- Even with calibration, RSSI data are still not accurate
- False negative: “Infected person RSSI r=3 meters, real distance 1.8 meters”
- False positive: “Infected person RSSI r=1.8 meters, real distance 2.5 meters”

BLE beacon ID anonymity







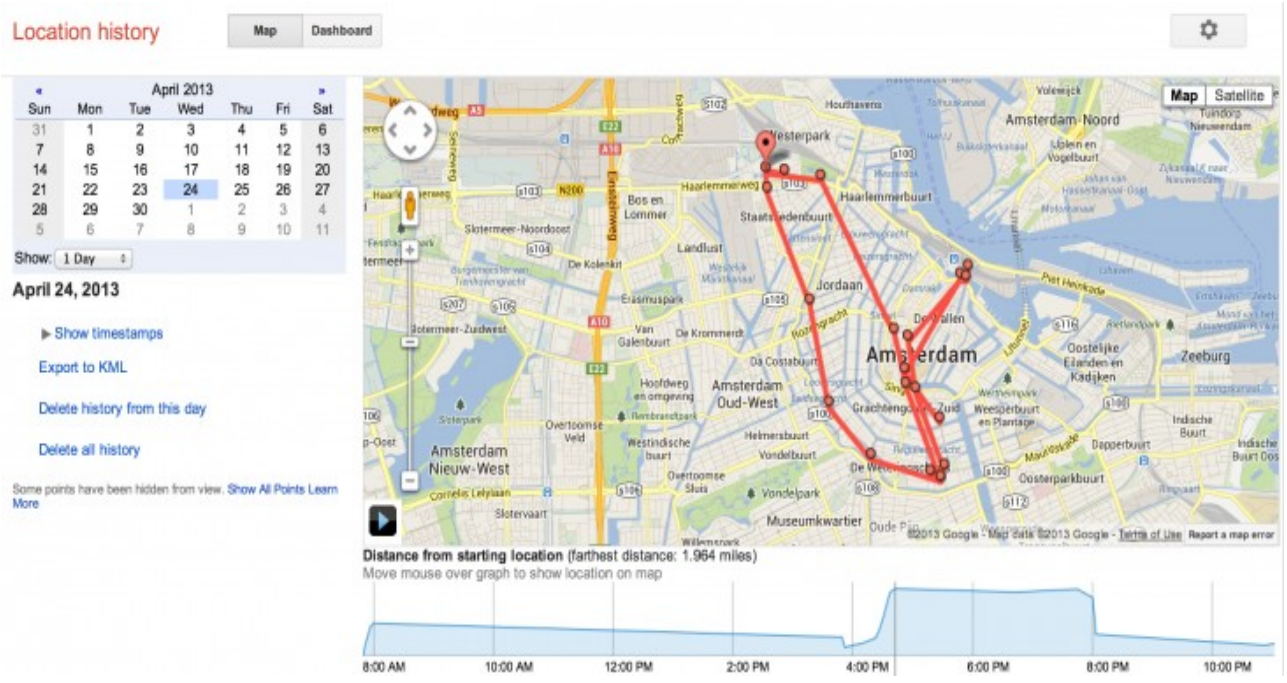
- The BLE beacon wraps around ephemeral IDs to more than one operation (contact tracing + wireless headset)
- OS controlled
- The wireless device data are ID-ed (Manufacturer ID)
- This forms the basis for an adversarial linkage attack
- Record traffic, isolate (visually) someone by distance (RSSI) that has a specific device whose Manufacturer ID you know and you have your person.

Bluesnarfing attacks

- Many older (> 3 year) smartphone are vulnerable
  - Android =< 9
  - IOS <= 11
- Results in loss of sensitive information (contact lists, SMS, phone digital contents)
- BLE LoS range increases the attack vector

	BLUETOOTH V2.1	BLUETOOTH 4.0 (LE)	BLUETOOTH 5 (LE)
Range	Up to 100 m	Up to 100 m	Up to 400 m
Max range (free field)	Around 100 m (class 2 outdoors)	Around 100 m (outdoors)	Around 1,000m (outdoors)
Frequency	2.402 – 2.481 GHz	2.402 – 2.481 GHz	2.402 - 2.481 GHz
Max data rate	1- 3 Mbit/s	1 Mbit/s	2 Mbit/s
Application Troughput	0.7-2.1 Mbit/s	Up to 305 kbit/s	Up to 1,360 kbit/s
Topologies	Point-to-point, scatternet	Point-to-point, mesh network	Point-to-point, mesh network

Privacy aspects



- GPS coordinates (A-GPS) are not accurate to the meter
- Authorities have embedded the GPS functionality to experiment with density maps
- GPS data are personal, should never be used as part of contact tracing solutions
- GPS info may also be misused

Privacy aspects (2)

- Are large amounts of (theoretically) anonymized health records sensitive data?
- How should they be stored and processed?
  - Standards
  - Audit facilities
  - Interfacing these data to scientists/health authorities

EU recommendations

Contact tracing solutions should:

- Limit usage of personal data
- Encrypt anonymous data
- Place time limit on data storage
- Offer data accuracy
- Offer data interoperability across the EU

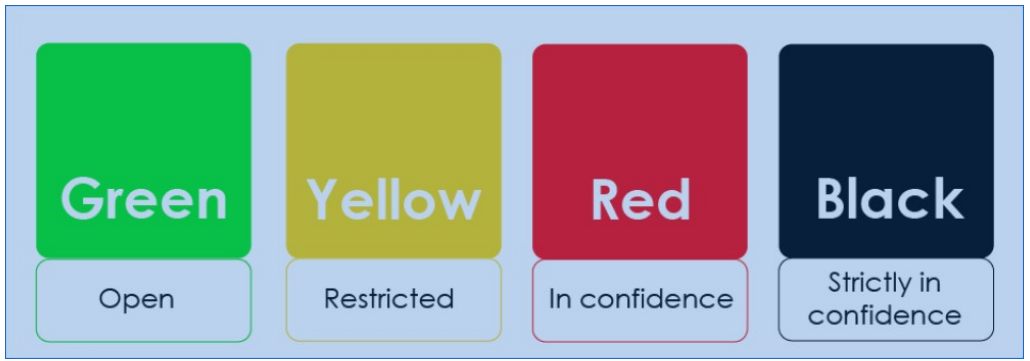
Source: An EU approach for efficient contact tracing

Genomic medicine as a paradigm

- Genomic medicine as a paradigm
- Large amounts of anonymous data
- Genomic data are also prone to linkage attacks
- Millions of manhours of infrastructure to ensure the data are shared
  - -Securely (authentication, availability)
  - -By relevant people (selection committees)
  - -In a transnational scalable manner
  - -Already complying to GDPR

Source: https://ega-archive.org/

University of Oslo data classification



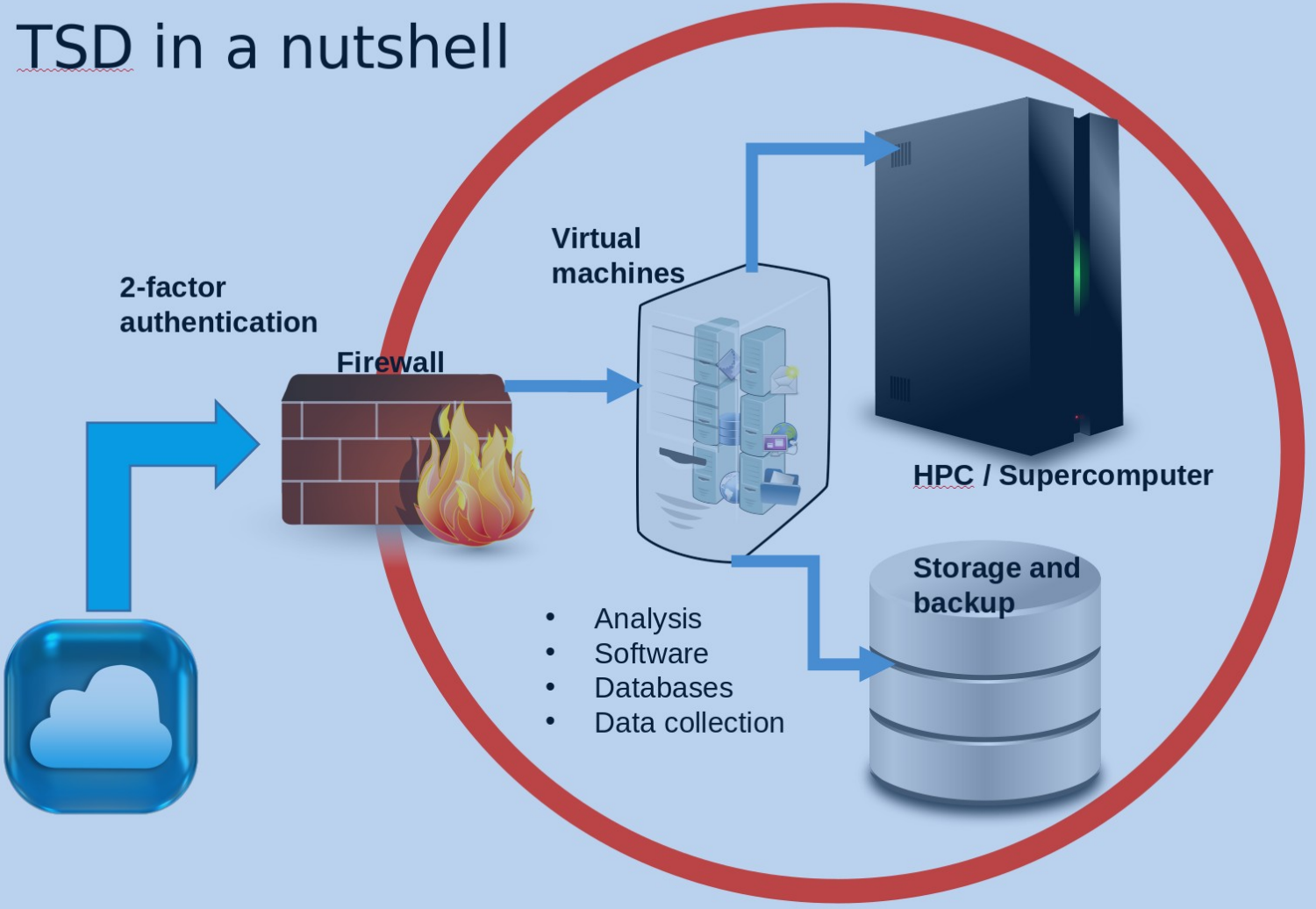
Examples of ‘Black’ ‘Strictly in confidence category:

- Large amounts of sensitive (patient identifiable) data
- Large amounts of anonymized data about people’s health
- Research data of huge economic value

Source: UiO data classification guide

UiO : University of Oslo

TSD in a nutshell



Conclusions

- Compare and contrast national government efforts to well established frameworks (EGA, TSD)
- Need for an end-to-end expert review:
  - Entire application stack (mobile OS, application, backend infrastructure) review
  - Precautions for data locality and private cloud usage
  - Proven BLE calibration standards

Questions

georgios@steelcyber.com