Protecting sensory data against sensitive inferences

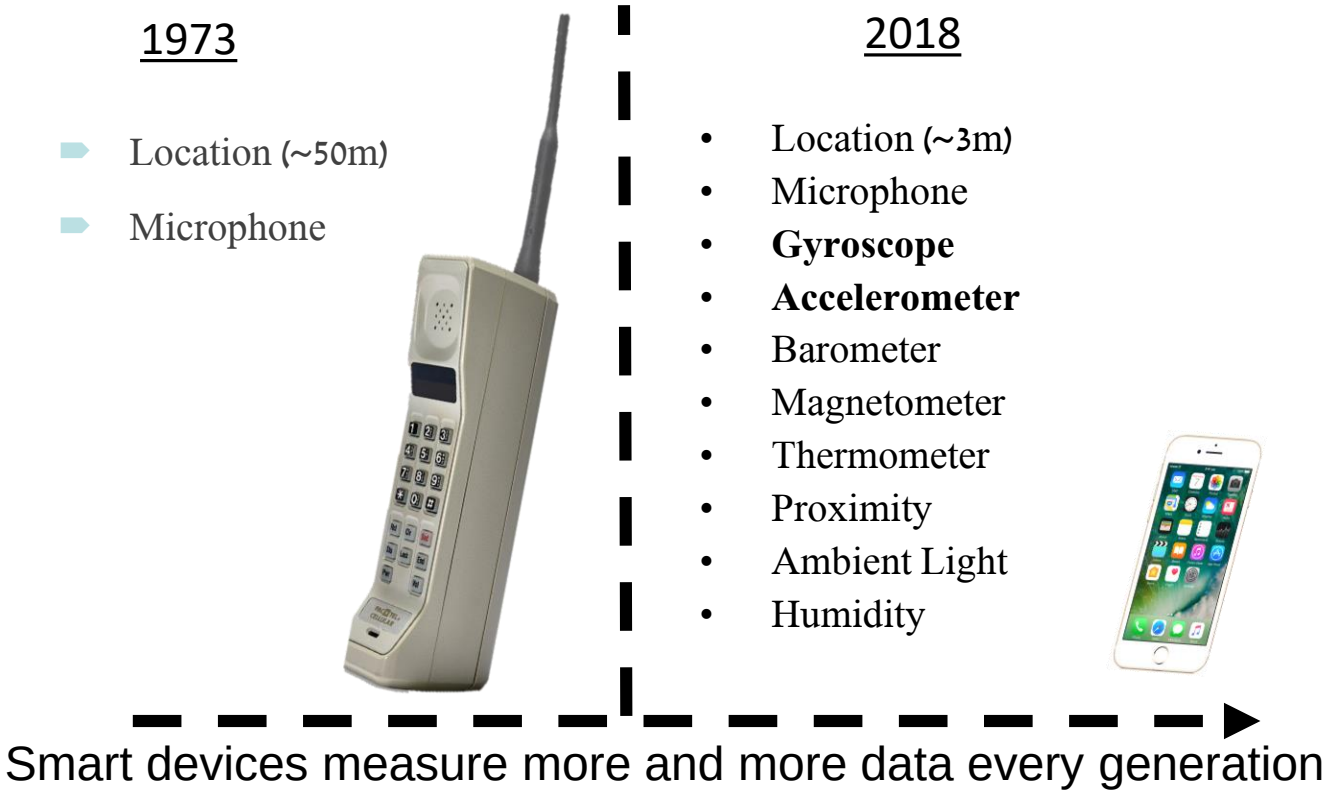
Mohammad Malekzadeh, Richard G. Clegg, Andrea Cavallaro, Hamed Haddadi

Published in: *Workshop on Privacy by Design in Distributed Systems (W-P2DS) 2018*

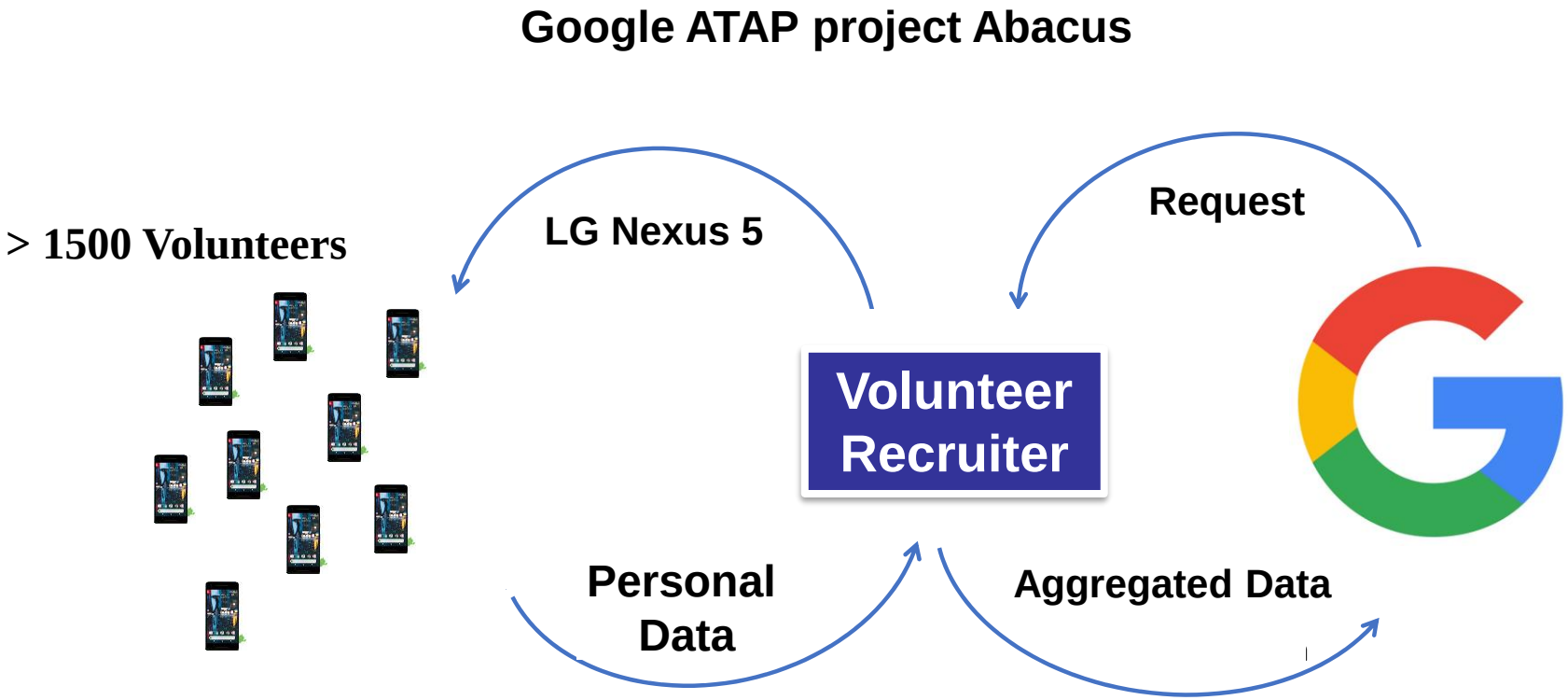
Centre for Intelligent Sensing

Queen Mary University of London

Context


1973
Location (~50m)
Microphone
2018
Location (~3m)
Microphone
Gyroscope
Accelerometer
Barometer
Magnetometer
Thermometer
Proximity
Ambient Light
Humidity
Smart devices measure more and more data every generation

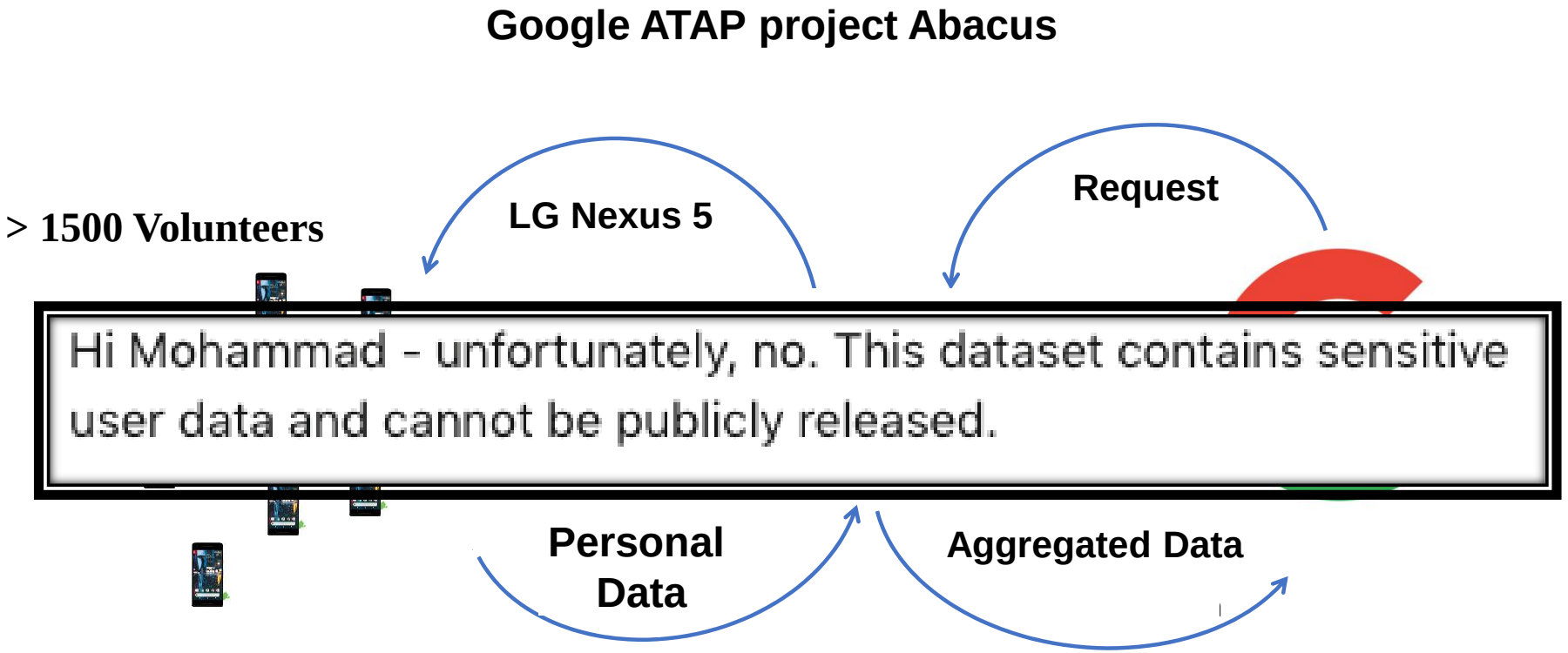
An example



- **Goal:** using biometric patterns, like motion, instead of password

Ref: DOI 10.1109/ACCESS.2016.2557846

An example



- **Goal:** using biometric patterns, like motion, instead of password

Ref: DOI 10.1109/ACCESS.2016.2557846

MotionSense Dataset

- Same Activity Set : 6 ADL activities
- Same Place
- Same Phone in the Front Pocket
- Accelerometer and Gyroscope
- **24 Different Subjects :**
  - **Gender:** 14 male - 10 female
  - **Age:** [18 – 40] years old
  - **Weight:** [45 ,105] kg
  - **Height:** [160 , 195] cm

DATABOX

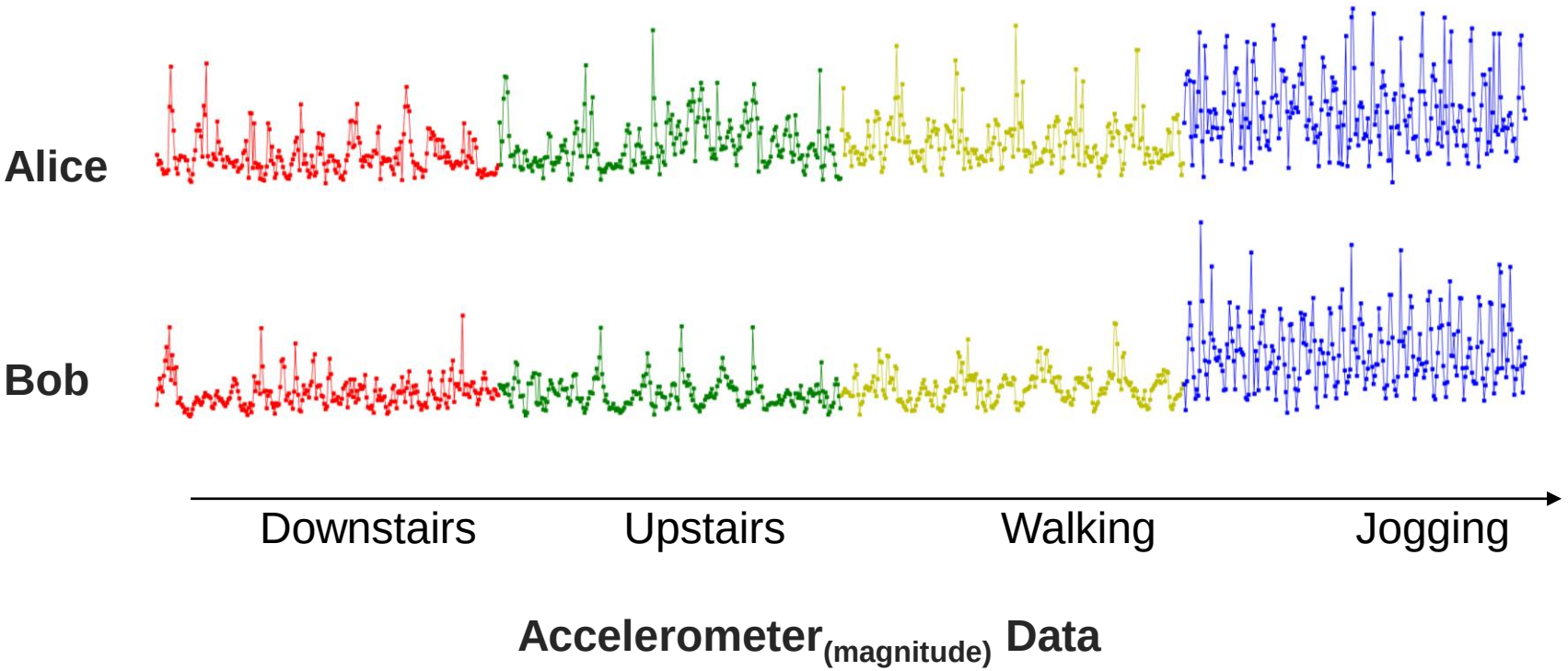
SensingKit

Queen Mary University of London

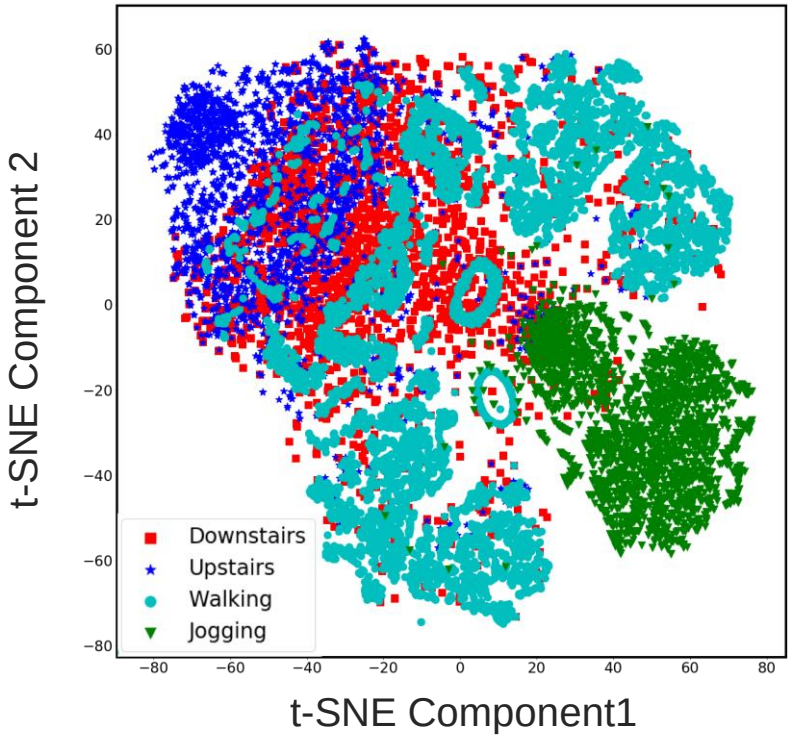
CIS centre for intelligent sensing

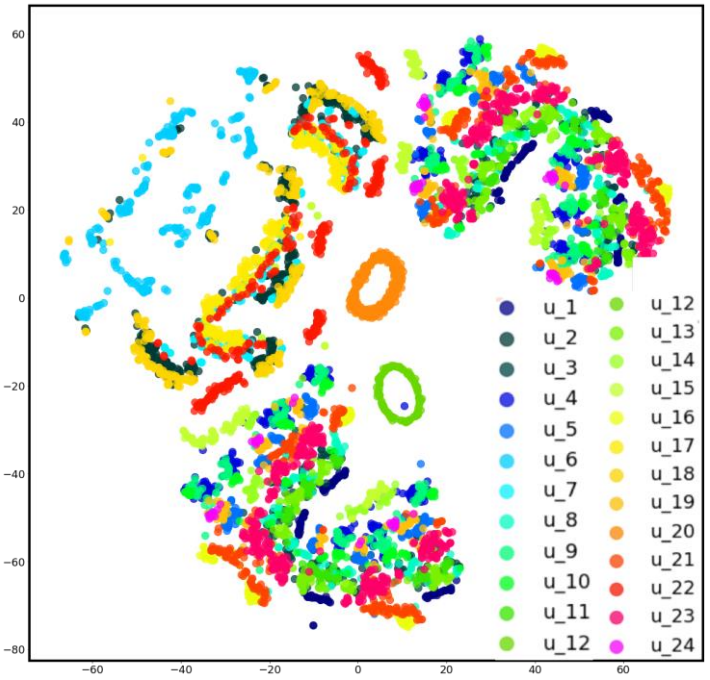
Queen Mary University of London

MotionSense Dataset

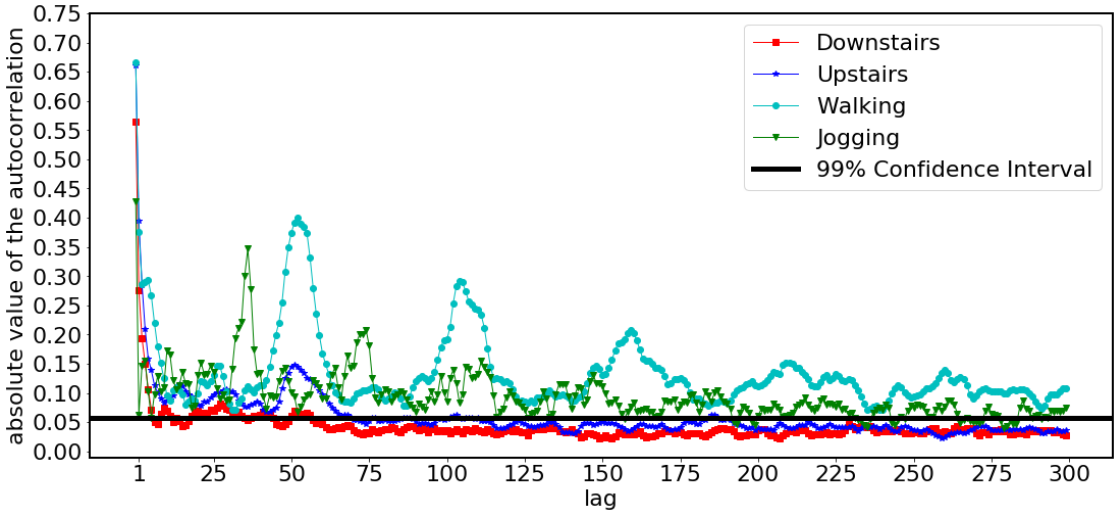


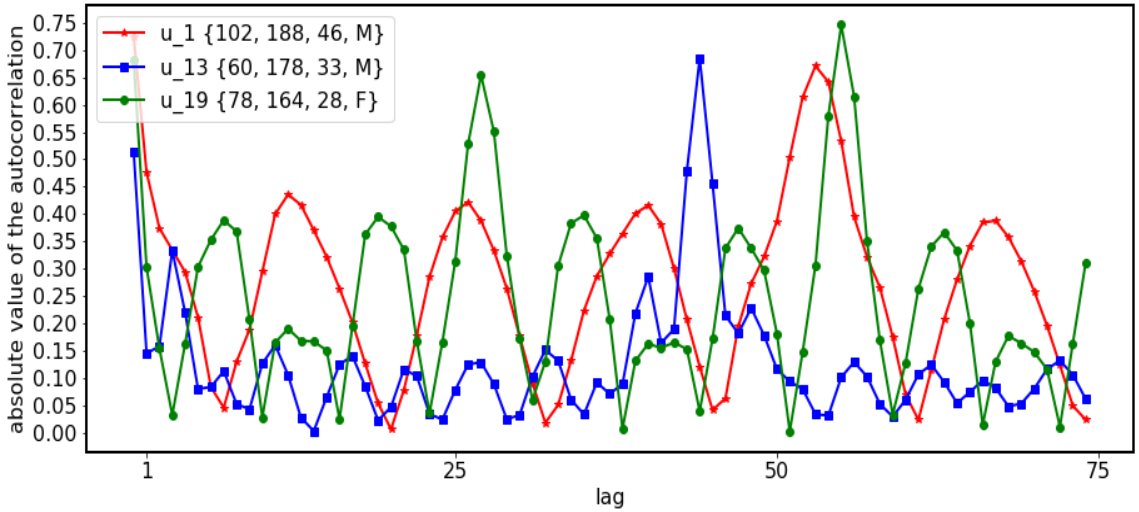
Visualisation





Autocorrelation



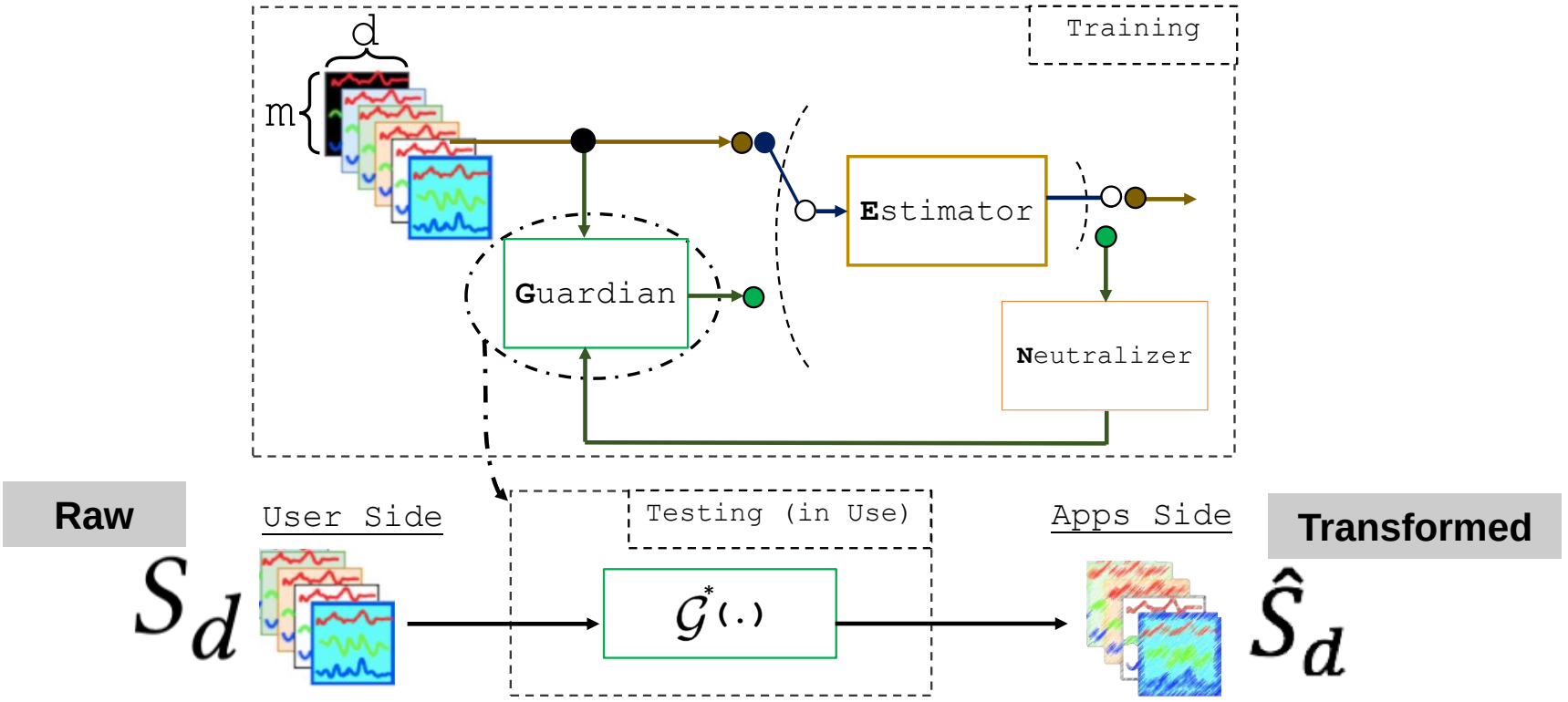


Classification

- 1-D Accelerometer$_{(magnitude)}$: (50Hz)
- Time-Window 5 second
- Deep Convolutional Network

	Classification Accuracy
activity	**~ 98%**
gender	**~ 96%**
Identity	**~ 89%**

Proposed Framework

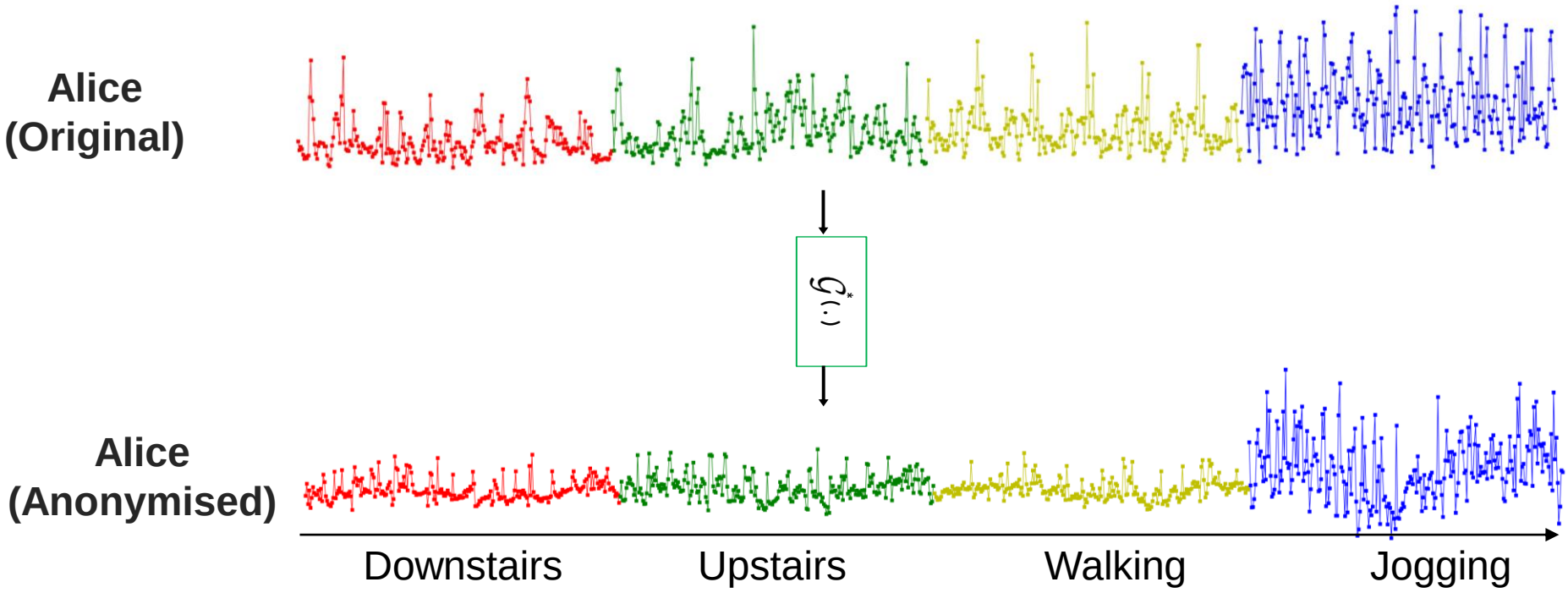


$$\mathcal{G}^*(.) = \underset{\mathcal{G}(.)\in\mathcal{F}}{\operatorname{argmin}} \left(p\left(I_s\left(\hat{S}_d\right)\right) - p\left(I_n\left(\hat{S}_d\right)\right) \right)$$

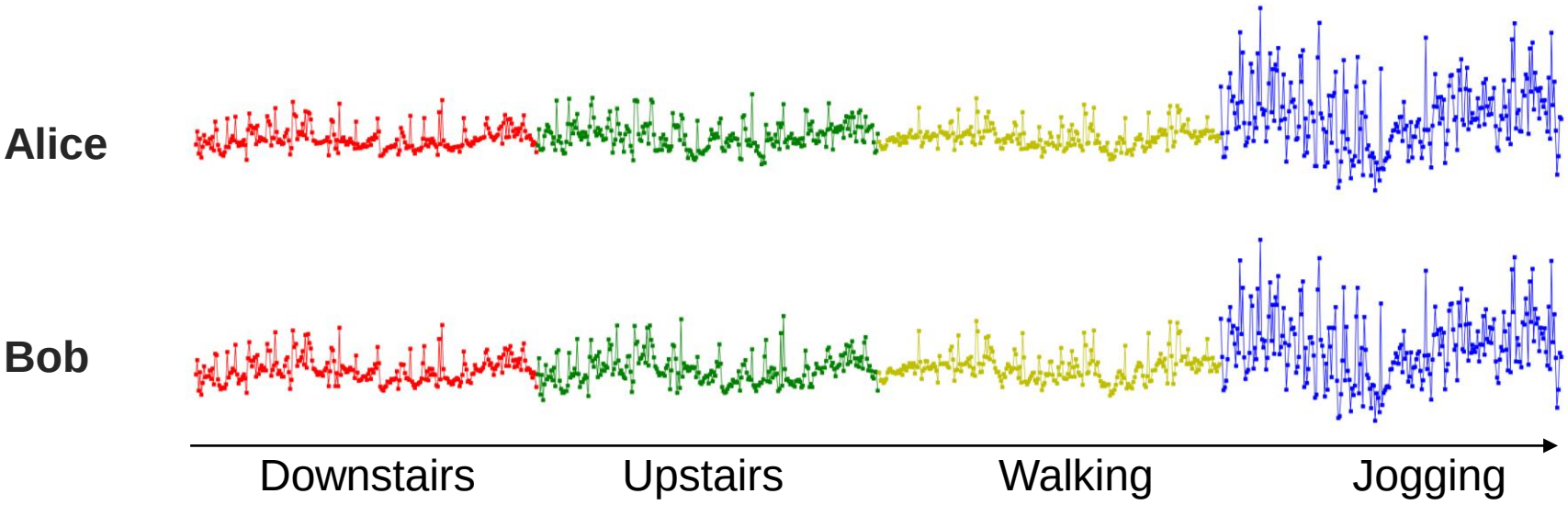
Sensitive Inferences

Non-Sensitive Inferences

After Transformation



Transformed Data

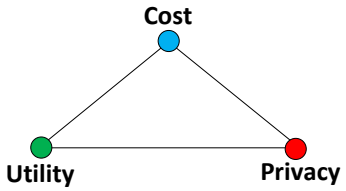


Results

			Classification Accuracy	
Setting	Dataset	Inf.	S_d	$\hat{S}_d$
Train: 2 out of 3	**MotionSense**	**activity**	95.08	93.71
		gender	95.15	49.32
	MobiAct	**activity**	94.31	90.46
		gender	93.74	49.83
Train: ¾ subjects	**MotionSense**	**activity**	86.33	85.19
		gender	75.35	52.16
	MobiAct	**activity**	70.49	65.01
		gender	66.18	45.54

Next Steps

- **Practical:**



  - The **Cost** of the solution on **Edge** devices?
  - Removing identifiable motion patterns.

- **Theoretical:**
  - Provide a **statistical guarantee** (probabilistic bound)
    - Differential Privacy : **Composition Theorem?**
    - Mutual Information : **Joint Distributions?**

Thanks!

Repository of the MotionSense Dataset:
bit.ly/eli-dw18