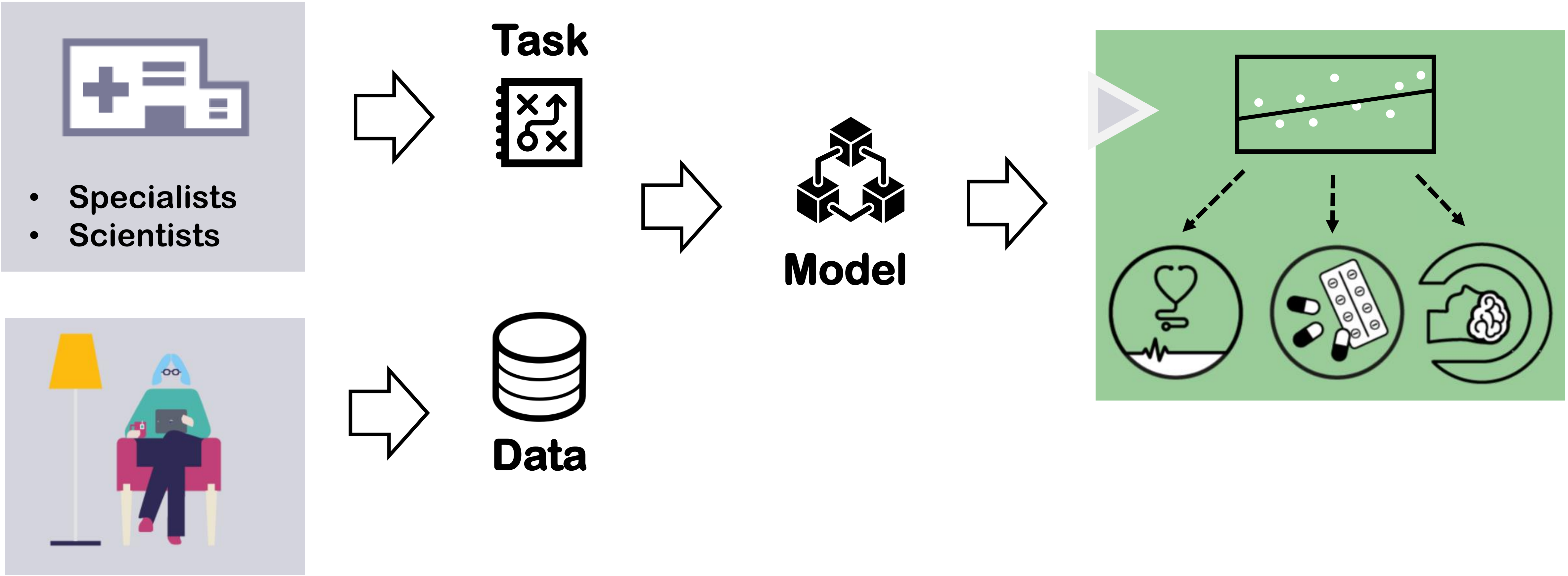
Differential Privacy Approaches in a Clinical Trial

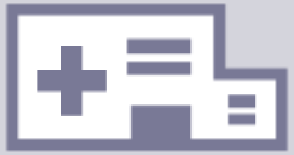
Martin Leuckert

Otto-von-Guericke University Magdeburg

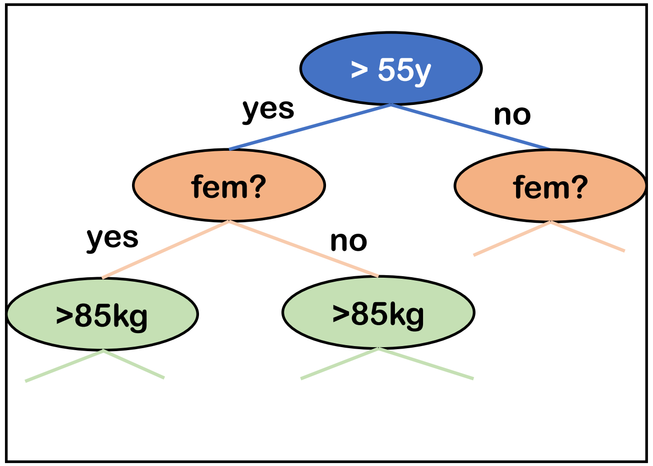
INTRODUCTION



INTRODUCTION



- Specialists
- Scientists



Private data

Age: 68
Weight: 78kg
Sex: Female
Diabetes Type: II

Sensitive data

Problem Statement

European Regulations:

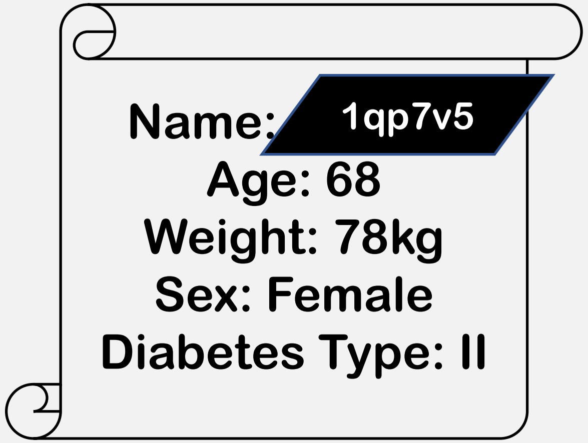
- General Protection Regulation (GDPR)
- RL 93/42/EWG (MEDDEV)
- GCP Directive (Directive 2005/28/EC)
- Clinical Trial Directive' (Directive 2001/20/EC)

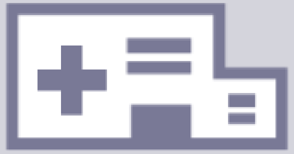
National Regulations

- Medizinproduktegesetz (MPG)
    - 90/385/EWG
    - 90/42/EWG
    - 98/79/EG

Problem Statement

So what exactly is the problem?

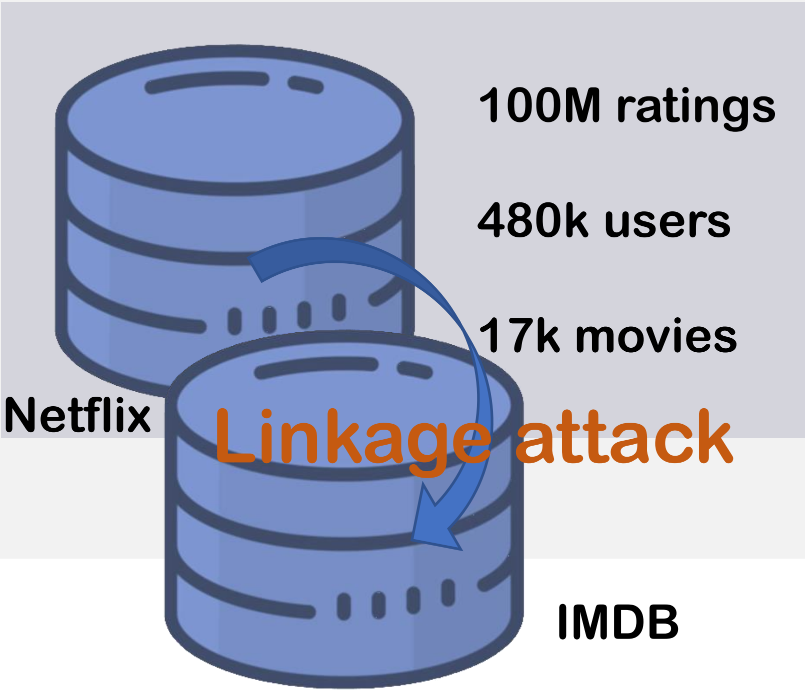




- Specialists
- Scientists

Hospitals anonymize or rather pseudonymize the data but is subject still at risk? YES.

NETFLIX PRIZE 2006



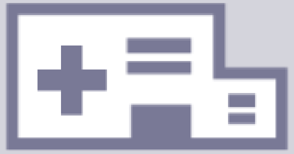
Anonymized movie ratings

1qp7v5	The Da Vinci Code	★★★
j8R04j	Inside Man	★★★★★
vyj6tG	Night at the Museum	★
0y6Hd	Ice Age: The Meltdown	★★★★
btWz3	Step Up	★★★★

Many Netflix users rated the same movies similarly at IMDB.
„Robust De-anonymization of Large Sparse Datasets“ by Arvind Narayanan and Vitaly Shmatikov, University of Texas at Austin, 2008

SO WHAT EXACTLY IS THE PROBLEM?

Name: 1qp7v5
Age: 68
Weight: 78kg
Sex: Female
Diabetes Type: II

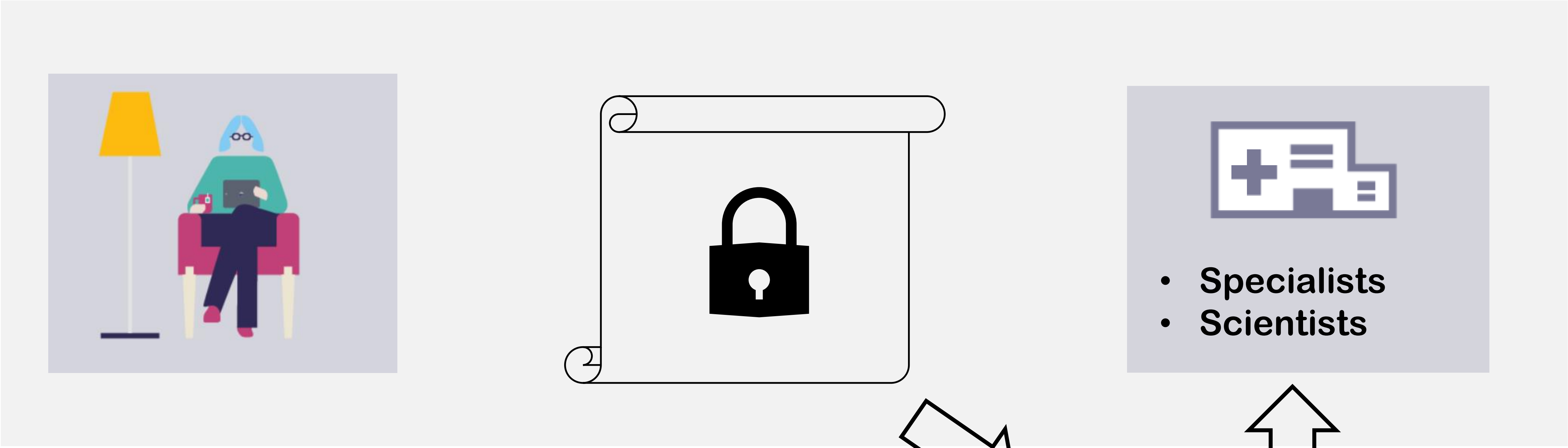


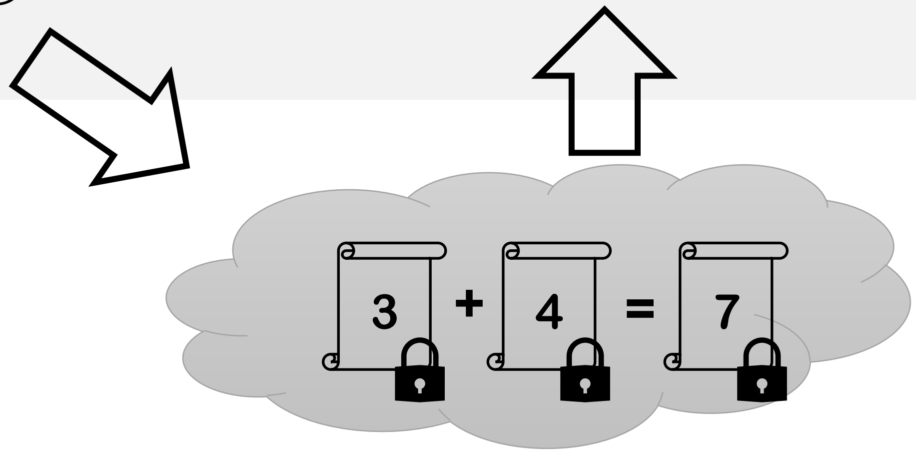
- Specialists
- Scientists

Third Party

INVOLVED THIRD PARTIES FOR DATA ANALYSIS MUST BE CONSIDERED MALICIOUS

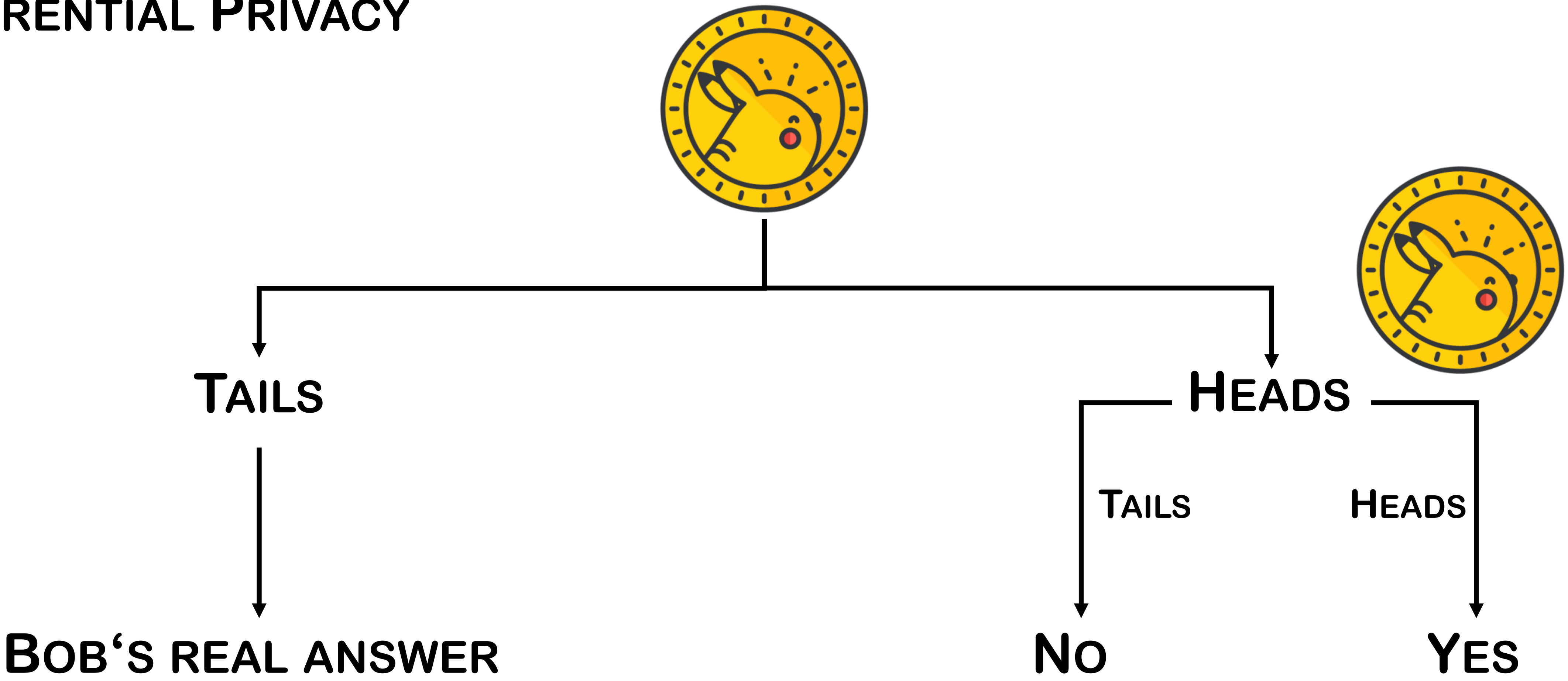
Pure crypto Solutions exist





Fully Homomorphic encryption
Garbled Circuits, secret sharing
Can be efficient for some specific cases

Differential Privacy



DIFFERENTIAL PRIVACY

„The Algorithmic Foundations of Differential Privacy“ by Cynthia Dwork, Microsoft Research & Aaron Roth, University of Pennsylvania, 2014.

“Differential privacy” describes a promise, made by a data holder, or curator, to a data subject: “You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available.”

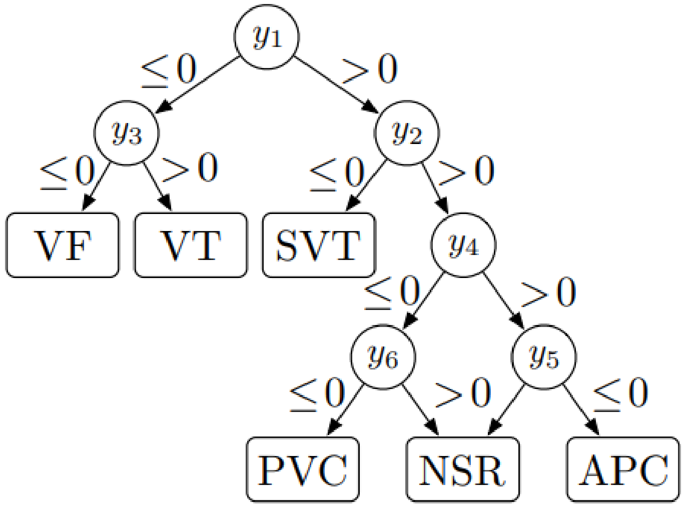
“Data cannot be Fully Anonymized and Remain Useful.”

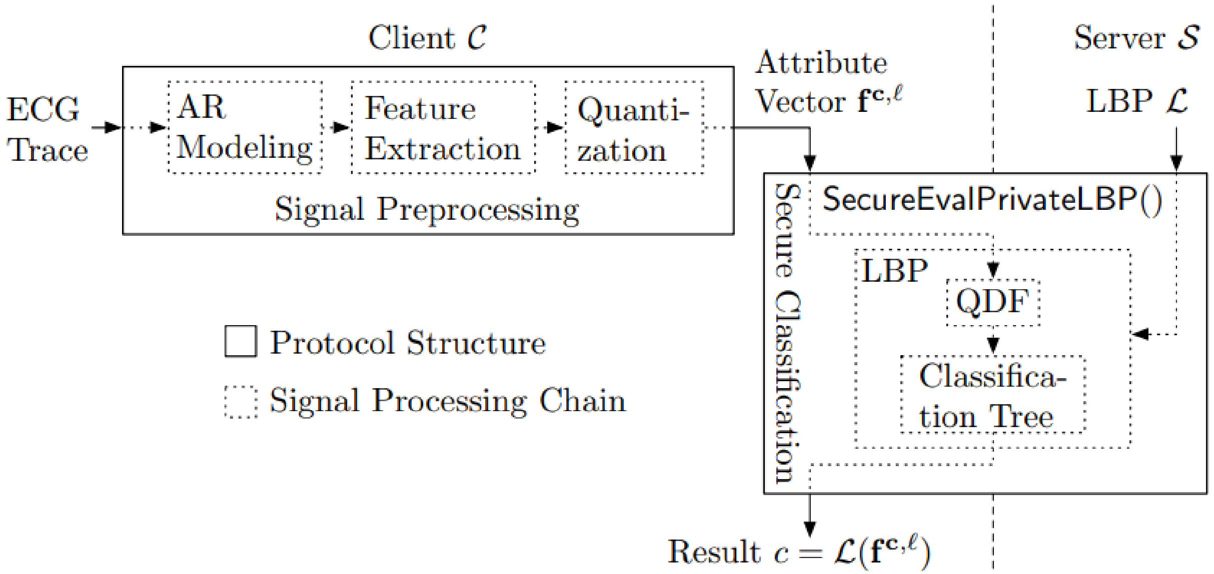
Can I call my algorithm $\mathcal{M}$ ε-DP?

$$L^{\xi} = ln\left(\frac{Pr[\mathcal{M}(x) = \xi]}{Pr[\mathcal{M}(y) = \xi]}\right)$$

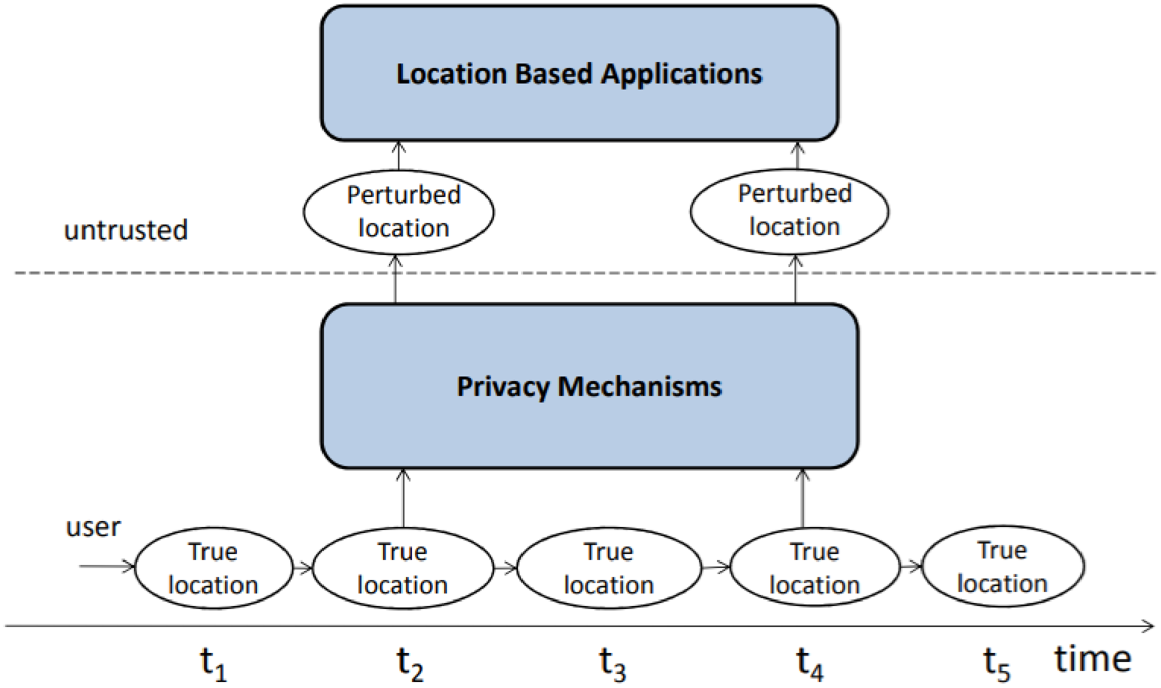
$\mathcal{M}$ **is called** ε**-DP, if and only if** $\left|L^{\xi}\right| \leq \varepsilon$

EXAMPLE 1: "EFFICIENT PRIVACY-PRESERVING CLASSIFICATION OF ECG SIGNALS", BARNI ET AL., 2009

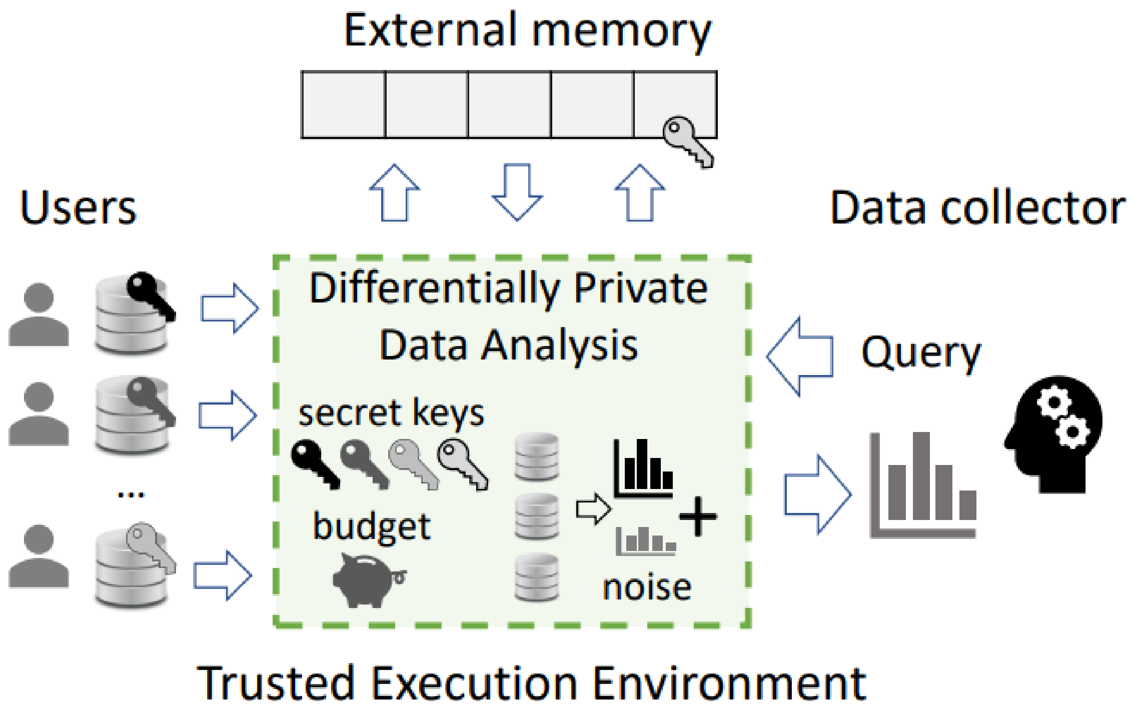




EXAMPLE 2 - DIFFERENTIAL PRIVACY: "PROTECTING LOCATIONS WITH DIFFERENTIAL PRIVACY UNDER TEMPORAL CORRELATIONS", YONGHUI, LI, EMORY UNIVERSITY AT ATLANTA, 2015.



EXAMPLE 3 - HARDWARE SOLUTION: "AN ALGORITHMIC FRAMEWORK FOR DIFFERENTIALLY PRIVATE DATA ANALYSIS ON TRUSTED PROCESSORS", ALLEN, OHRIMENKO ET AL.

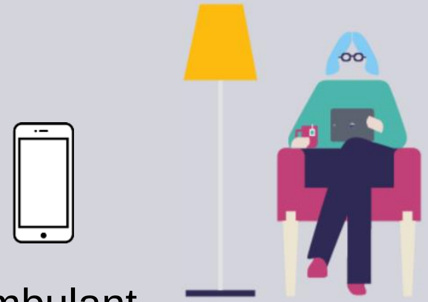


SMART PREVENT DIABETIC FEET: 300 SUBJECTS (150 USING DEVICE)

ALL

Classic onsite screening every six months

STUDY GROUP



Ambulant

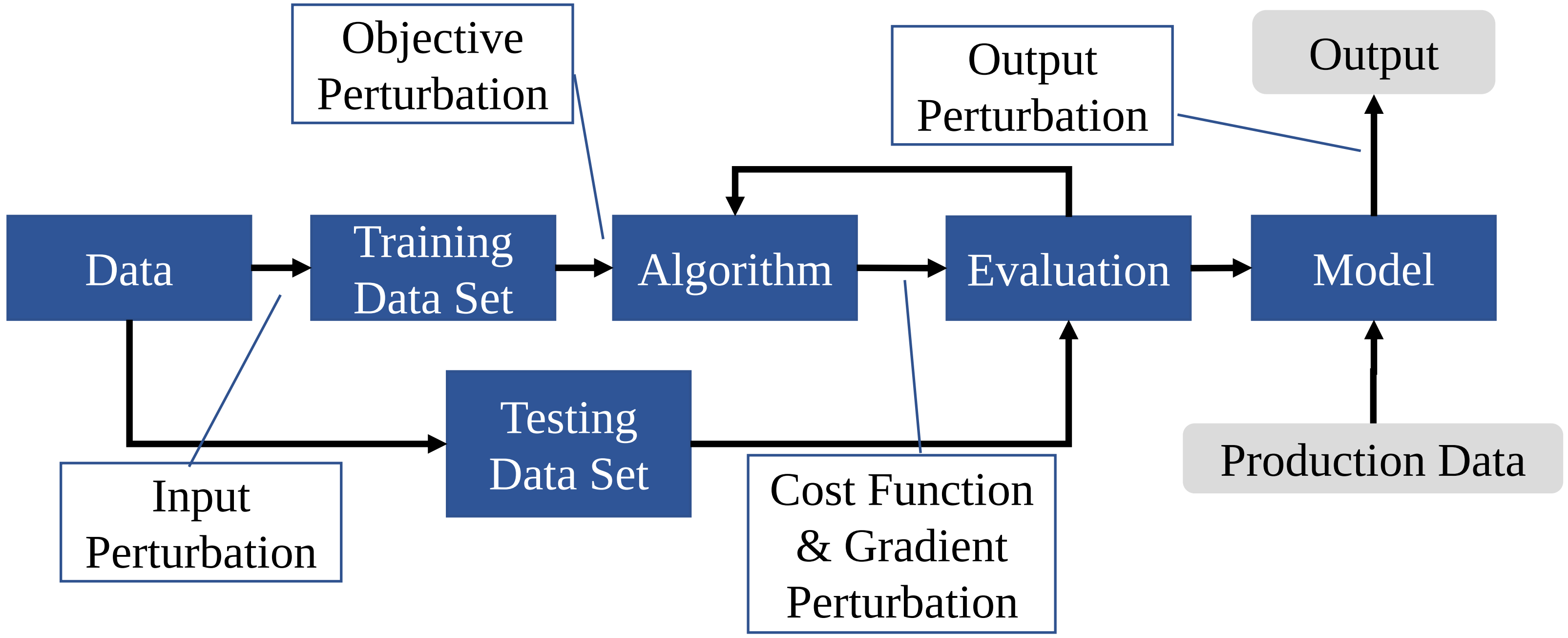
Data inspection

Combining results for a more sophisticated treatment of the subjects

Research Design & Methods

Compare two basic mechanisms for numeric input perturbation

- Laplacian Mechanism
- Functional Mechanism



Research Design & Methods

Experimental Setup

Hardware & Environment

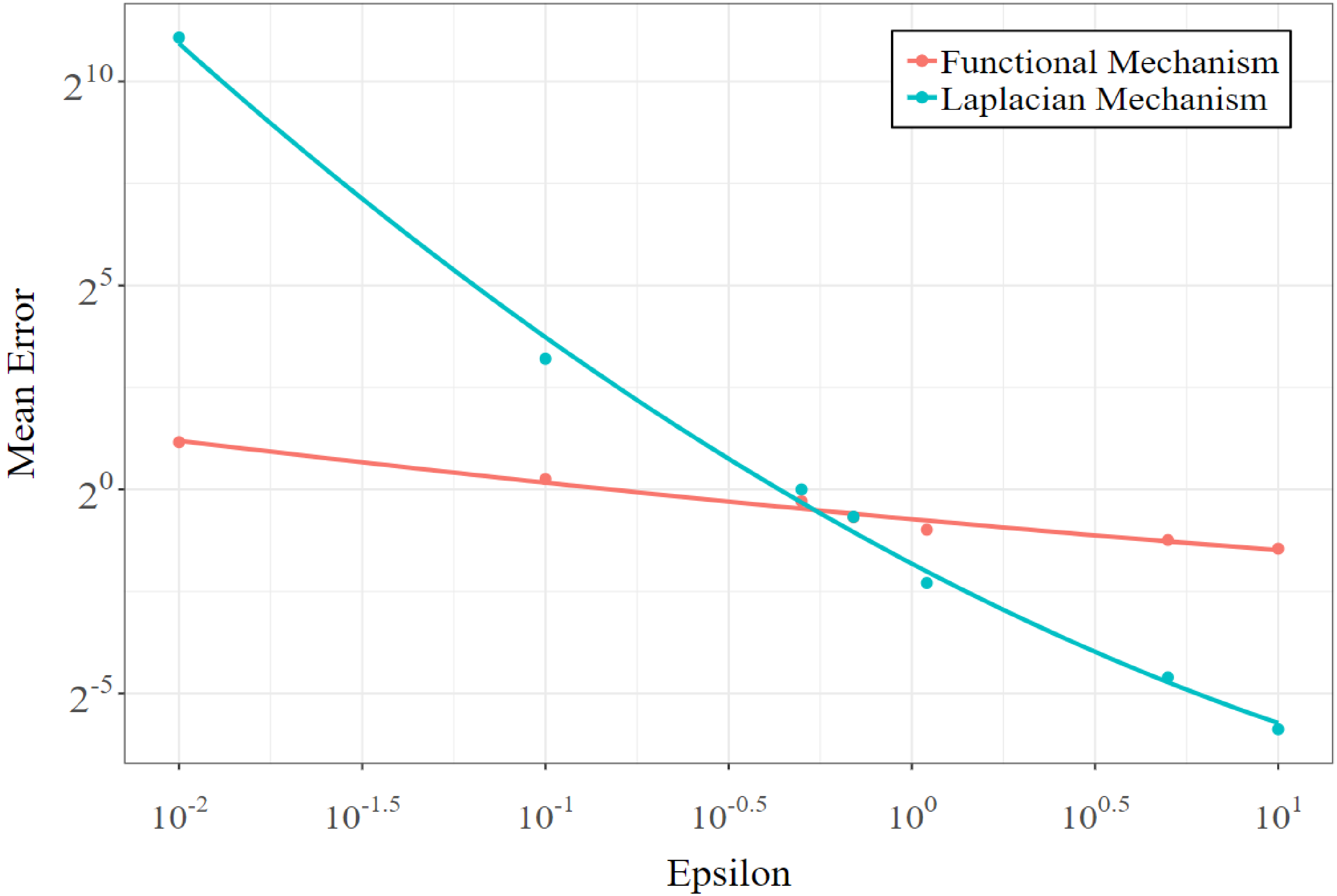
- Intel Core i7-8665U
- 48GB RAM
- MS ML.NET

Data

- 10-fold cross-validation
- 200k data records

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - y_i^*\right)^2$$

OVERVIEW



(Privacy loss)

Conclusion

We explored Differential Private Machine Learning

Performance

Significantly faster than purely cryptographic Solutions

Usability

Increases with higher ε

Strongly influenced by chosen mechanism

Security Guarantees

Decrease with higher ε

Strongly influenced by chosen mechanism

CONCLUSION

DIFFERENTIAL PRIVACY CAN BE APPLIED TO CLINICAL TRIALS

BUT SOPHISTICATED DECISIONS REQUIRED ABOUT MECHANISM, PERTURBATION, ε, USABILITY

Differential Privacy can provide security and privacy, if applied correctly.

Out-of-the-box solutions required for smaller scale clinical trials.

Thank you.