DIGITAL IDENTITY  
PRIVACY THREATS AND BUSINESS OPPORTUNITY,  
WILL THE TECHNOLOGY FIX EVERYTHING?  

Horizon 2020 – PoSeID-on Public Launch Event  
Melek Önen  
PAPAYA (PlAtform for PrivAcY preserving data Analytics)  
Rome, October 25th 2018
DATA ANALYTICS AS A SERVICE (DAAAS)

Sensitive data: health data, web browsing, customer habits, identity, etc.
“Identity analytics [...] provide insights for making better IAM decisions. They enable organizations to take a contextual, dynamic, risk-based approach to IAM. The organization can bridge the gap between administrative controls and runtime access, add context and risk awareness to access decisions, and continuously monitor, detect, and remediate malicious behavior.”

**IDENTITY ANALYTICS - DATA ANALYTICS FOR DIGITAL IDENTITY**

**BENEFITS**

- Data presentation and visualization
- Identity correlation and profiling
- Behavioral and data analysis
- Risk scoring, computation and analysis
- Monitoring and alerting
“Identity analytics [...] provide insights for making better IAM decisions. risk-based approach to IAM. The organization can bridge the gap between administrative controls and runtime access, add context and risk awareness to access decisions, and continuously monitor, detect, and remediate malicious behavior.”

**BENEFITS**

- Risk scoring, computation, and analysis
- Identity correlation and profiling
- Behavioral and data analysis
- Monitoring and alerting
- Data presentation and visualization

DATA BREACHES IN 2018

Average Cost
Global: 3.86M$, Per record: 148$

Top 3 sectors
Health, Financial, Services

Factors increasing cost
Extensive migration to cloud
Third party involvement
Compliance failures

Factors decreasing cost
Extensive use of encryption
Use of security analytics
Provision of ID protection
GDPR EFFECT

Effective since May 2018

Fines ~ 20 million euros or 4% of turnover
PRIVACY REQUIREMENTS

WHAT TO PROTECT?

- Data
- Query
- Model

FROM WHOM TO PROTECT?

- Cloud server
- Querier
- Data collector

HOW TO PROTECT?

- Anonymisation
- Encryption
- User Control
PAPAYA – PLATFORM FOR PRIVACY PRESERVING DATA ANALYTICS

OBJECTIVES

- Privacy by design
- Different settings
- Integrated platform
TWO USE CASES – 4 USAGE SCENARIOS

Use Case 1: E-Health

- Single Data Owner
- Collaborative Data Analytics

Use Case 2: Web & Mobile

- Single Data Owner
- Third Party Querier
- Multiple Data Sources
- Third Party Querier

Funded by Horizon 2020 Framework Programme of the European Union
PAPAYA PETS MODULE – EXAMPLE

Compute \(NN(x)\), without knowing \(x\) (Prediction)

Data protection tools

\[ y = NN(x) \]

FHE

Functional Encryption (FE)

F

TEE

Secure Multiparty Computation (SMC)

\(x\)
NEURAL NETWORKS OVER PROTECTED DATA - CHALLENGES

Neural Network Layers

- Convolution, Fully Connected: linear functions
- Activation: sigmoid, tanh
- Pooling layer: softmax

Challenges

- Large data, non linear operations

PAPAYA Solution with ECG data

Approximation of non linear operations

- Privacy with efficiency
- Privacy with accuracy
PAPAYA CHALLENGES - SUMMARY

Data analytics
Sum, Average, Neural Networks, Clustering, etc.

Cryptographic Techniques
Homomorphic encryption,
Secure Multi-party Computation,
Secure data aggregation

Privacy control
What to protect? From whom to protect?

Dedicated cryptographic constructions & data analytics
User Dashboard
THANK YOU
FOR YOUR ATTENTION

and enjoy the conference

www.papaya-project.eu  @ProjectPapaya  PAPAYA Project  papaya@eurecom.fr

The research leading to these results has received funding from the European Union’s Horizon 2020 Research and Innovation Programme, through the PAPAYA project, under Grant Agreement No. 786767. This flyer reflects the view of the consortium only. The Research Executive Agency is not responsible for any use that may be made of the information it contains.