Secure and Privacy-Preserving Benchmarking for Artificial Intelligence in Health

Jean Louis Raisaro and Jean-Pierre Hubaux

ITU/WHO workshop on "Artificial Intelligence for Health"
November 14th , 2018, New York City, NY, USA

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

CHUV

DPPH
Data Protection in Personalized Health

Dr. Jean Louis Raisaro

- Medical informatics and data protection specialist in the Precision Medicine Unit at Lausanne University Hospital (CHUV), Switzerland.
- 2012 - 2018: PhD in Computer Science (privacy and security) at EPFL
  - Thesis: “Privacy-Enhancing Technologies for Medical and Genomic Data: From Theory to Practice”
- 2006 - 2012: BS + MS in Medical Informatics and Bioinformatics
  - BS + MS in Medical Informatics and Bioinformatics at University of Pavia, Italy
  - Thesis: “An Automatic SNOMED CT Encoder for Clinical Free-Text”

Growing Concern: Medical Data Breaches

Around 5 declared breaches per week, each affecting 500+ people

https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf

Under Investigation | Archive | Help for Consumers

As required by section 13402(e)(4) of the HITECH Act, the Secretary must post a list of breaches of unsecured protected health information affecting 500 or more individuals. The following breaches have been reported to the Secretary:

Cases Currently Under Investigation

This page lists all breaches reported within the last 24 months that are currently under investigation by the Office for Civil Rights.

Show Advanced Options

Breach Report Results

Expand All	Name of Covered Entity	State	Covered Entity Type	Individuals Affected	Breach Submission Date	Type of Breach	Location of Breached Information
	Ohio Living	OH	Healthcare Provider	6510	09/07/2018	Hacking/IT Incident	Email
	Rockdale Blackhawk, LLC d/b/a Little River Healthcare	TX	Healthcare Provider	1494	09/07/2018	Unauthorized Access/Disclosure	Electronic Medical Record, Other
	J.A. Stokes Ltd.	NV	Healthcare Provider	3200	09/05/2018	Hacking/IT Incident	Desktop Computer, Electronic Medical Record, Network Server
	Reliable Respiratory	MA	Healthcare Provider	21311	09/01/2018	Hacking/IT Incident	Email
	Port City Operating Company doing business as St. Joseph's Medical Center	CA	Healthcare Provider	4984	08/31/2018	Loss	Other Portable Electronic Device
	Carpenters Benefit Funds of Philadelphia	PA	Health Plan	20015	08/31/2018	Hacking/IT Incident	Email

DPPH – Data protection in personalized health

- 5 research groups across the ETH domain + SDSC (Swiss Data Science Center)
- Funding: 3 Millions CHFrs
- Duration: 3 years (4/2018 - 3/2021)
- Funding Program: ETH PHRT (Personalized Health and Related Technologies)

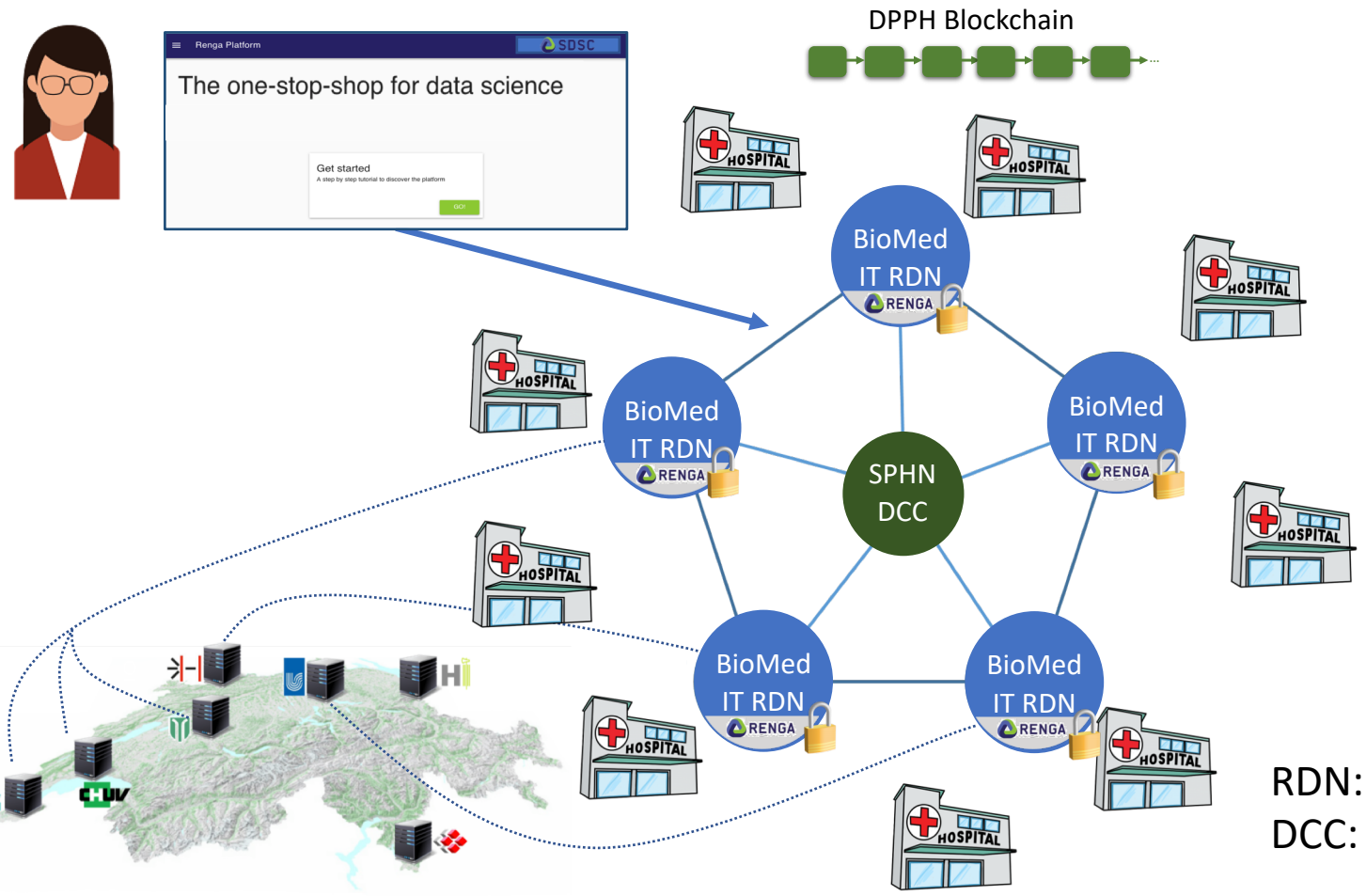
Strategic Focus Area
Personalized Health and Related Technologies

Project goals:

- Address the main **privacy, security, scalability, and ethical challenges** of data sharing for enabling effective P4 medicine
- Define an optimal **balance between usability, scalability and data protection**
- Deploy an appropriate set of **computing tools**

DPPH's Long-Term Vision

A One-Stop Shop for Collaborative Research on Health Data in the Context of Swiss Personalized Health Network



Platform requirements

- Interoperability (workflow and data)
- Reproducible research
- Big data scalability
- Auditability and Traceability
- Distributed data
- Secure data access
- Data protection compliance
- Privacy-conscious processing

RDN: Regional Data Node
DCC: Data Coordination Center

Technologies for Privacy and Security Protection

Traditional Encryption

- Protects data at rest and in transit
- Cannot protect computation

Homomorphic Encryption

- Protects computation in untrusted environments
- Limited versatility vs efficiency

Secure Multiparty Computation

- Protects computation in distributed environments
- High communication overhead

Trusted Execution Environments

- Protects computation with Hardware Trusted Element
- Requires trust in the manufacturer, vulnerable to side-channels

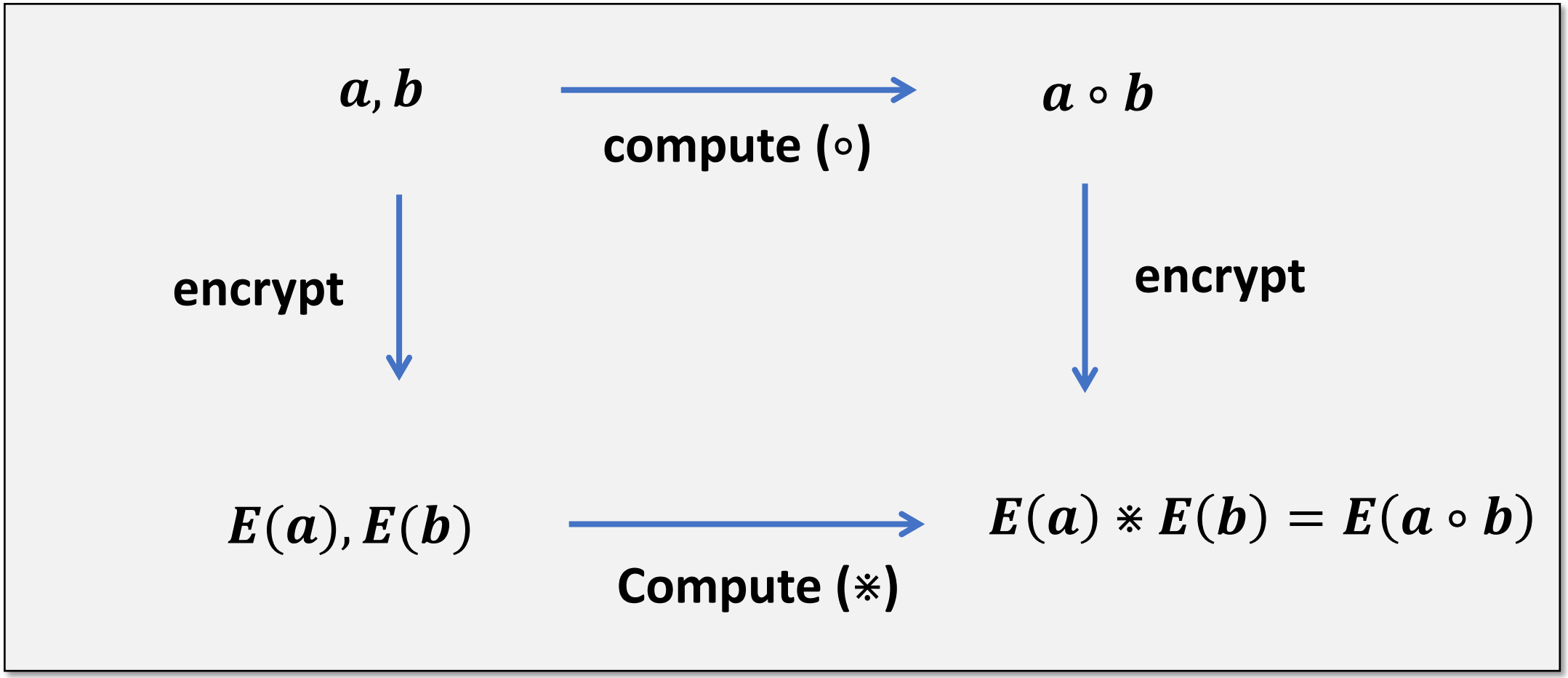
Differential Privacy

- Protects released data from inferences
- Degrades data utility (privacy-utility tradeoff)

Distributed Ledger Technologies (Blockchains)

- Strong accountability and traceability in distributed environments
- No privacy by default

Homomorphic Encryption

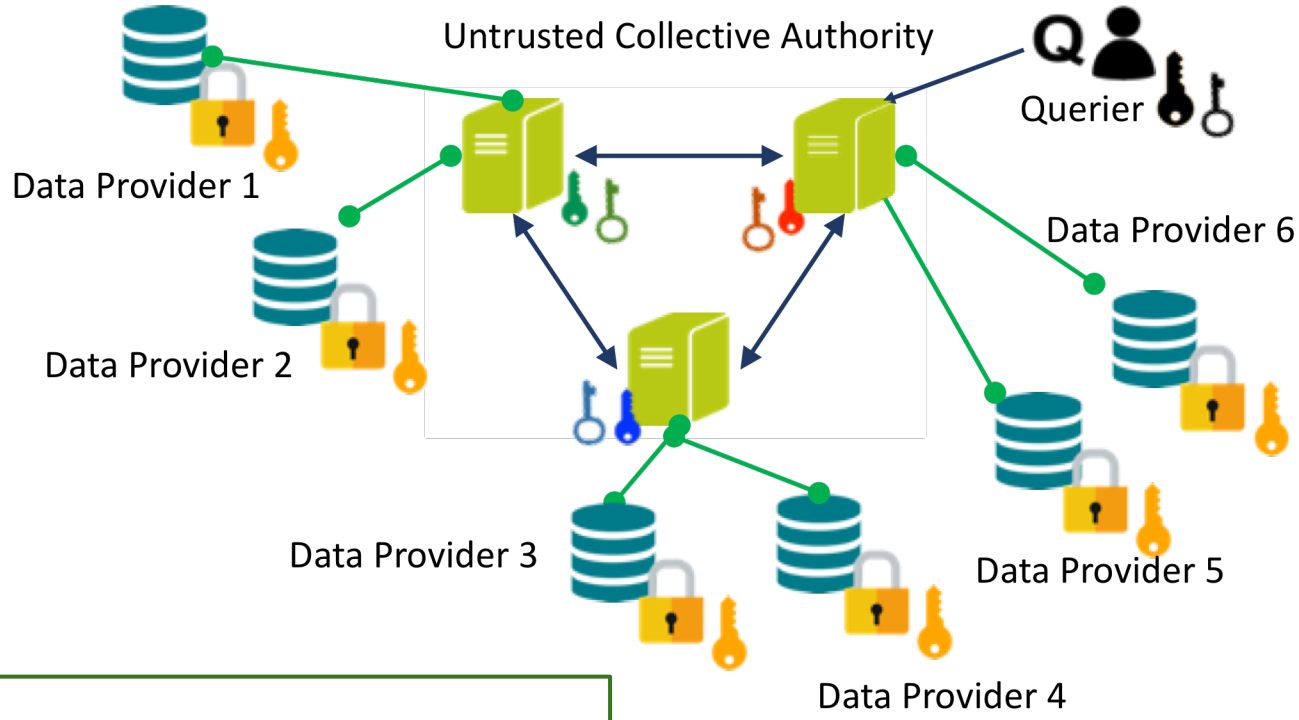


Homomorphic encryption enables computations directly on encrypted data.

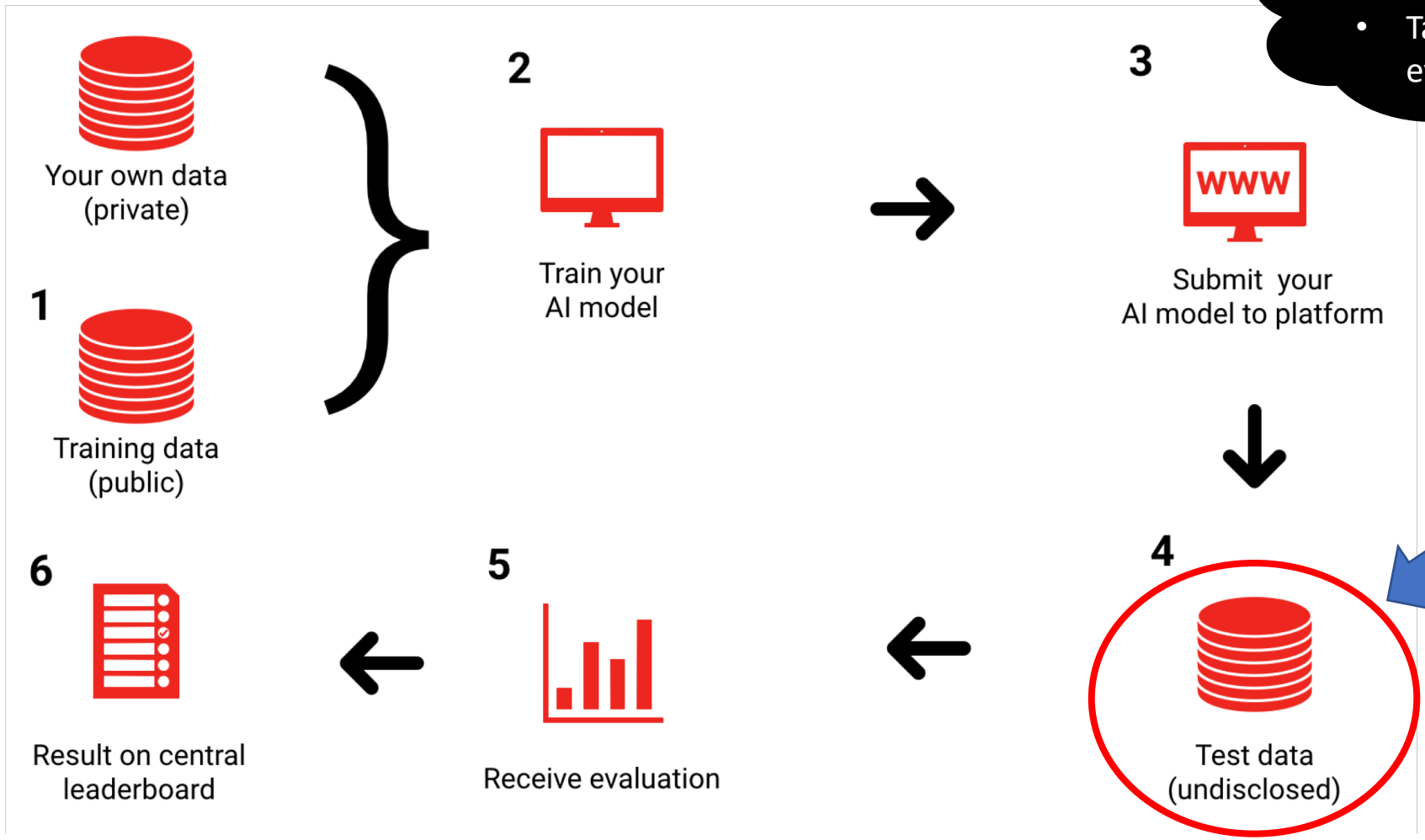
UnLynx: Framework for Privacy-Conscious Data Sharing

[Froelicher et al. 2017], [Raisaro et al. 2018]

- Trust is shared across a group of servers forming a **collective authority**
- They collaborate together to generate a **collective encryption key**
- The collective encryption key is used to encrypt the data and **can be compromised only if all servers are compromised**



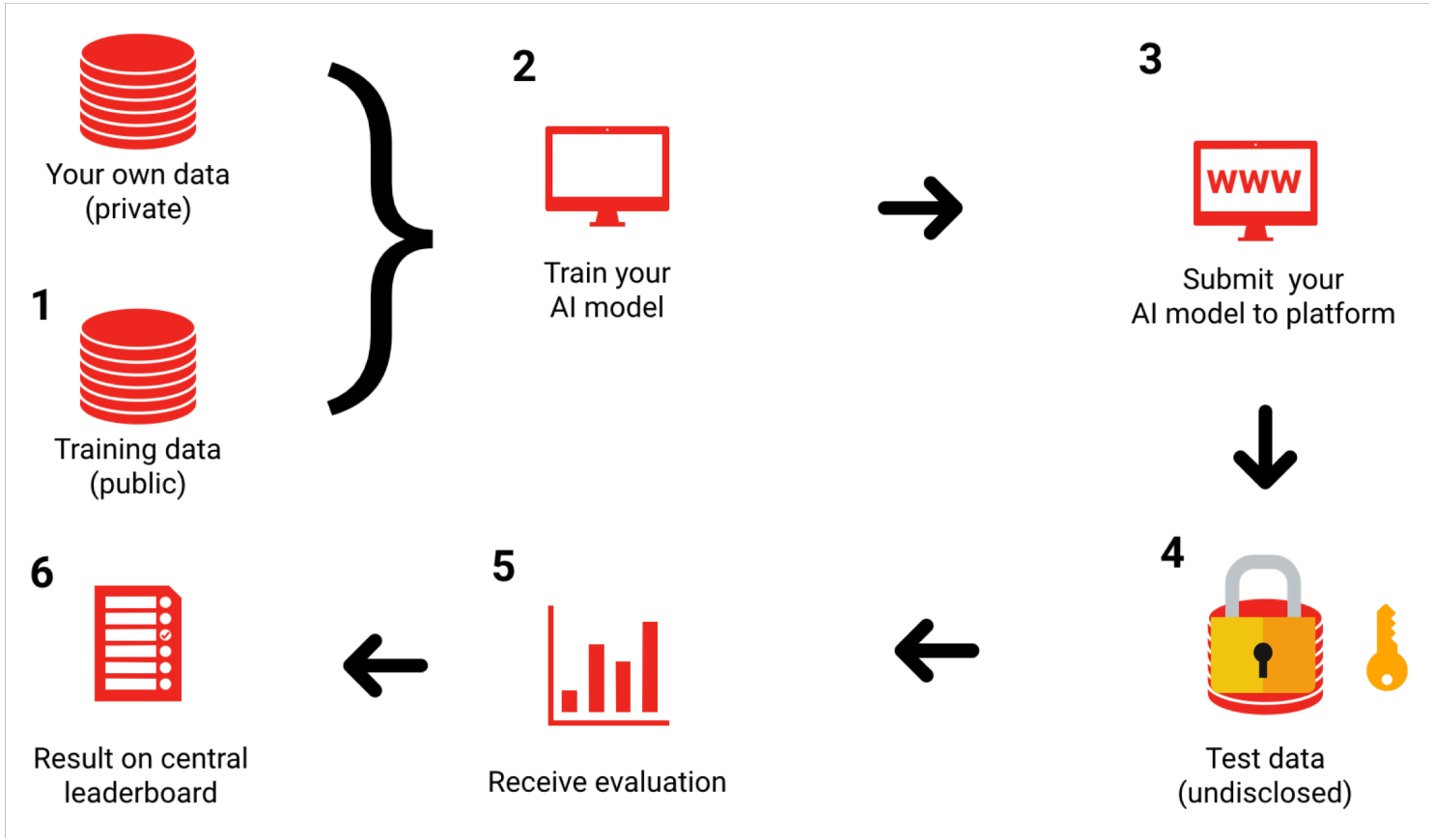
Benchmarking Pipeline

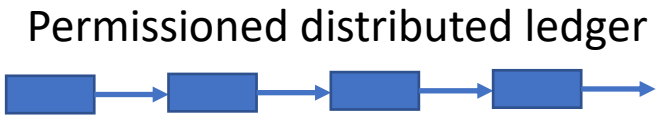


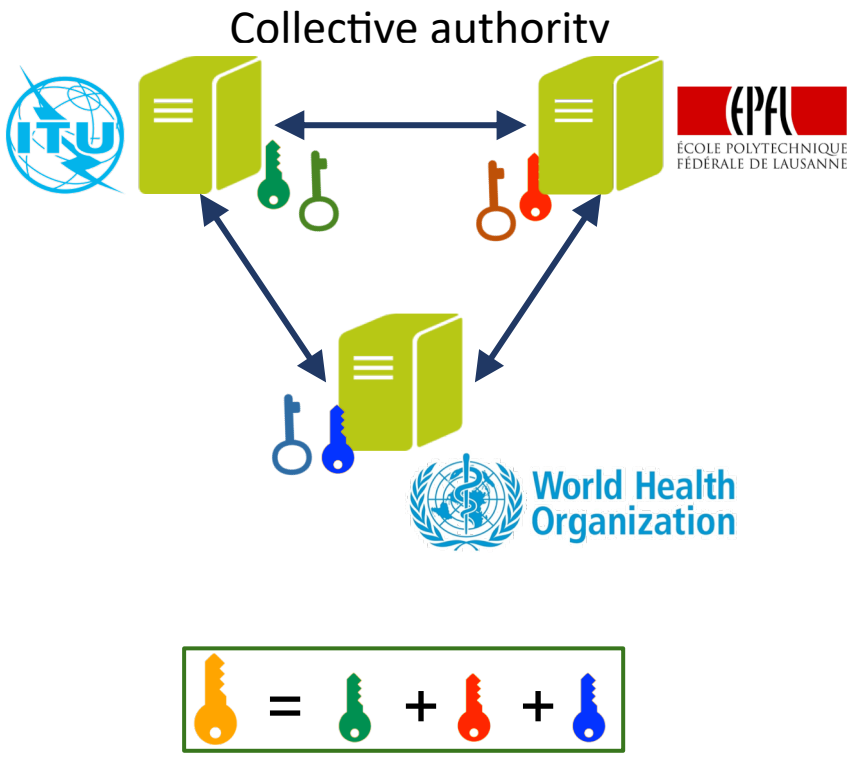
- Steal data in order to take an unfair advantage
- Tamper with the evaluation protocol

[1] Salathé M, Wiegand T, Wenzel M and Kishnamurthy R, *Focus Group on Artificial Intelligence for Health*, White paper https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FG- AI4H_Whitepaper.pdf

Privacy-Preserving Approach







- **Trust is distributed** within the collective authority
- Test data confidentiality is protected with **collective homomorphic encryption**
- Accountability is provided by the use of **permissioned distributed ledger** where all actions are immutably logged

Conclusion and Next Steps

- Worldwide, the confidentiality of health data is **in jeopardy**
- Standardization and regulation of AI in health can only be achieved if people trust the whole process to be **safe, secure and fair**
- Advanced privacy-enhancing technologies can be **effective enablers**

Next Steps:

- Explore the feasibility of integrating collective homomorphic encryption into an existing AI benchmarking platform (e.g., https://www.crowdai.org/)
- Develop a first proof of concept

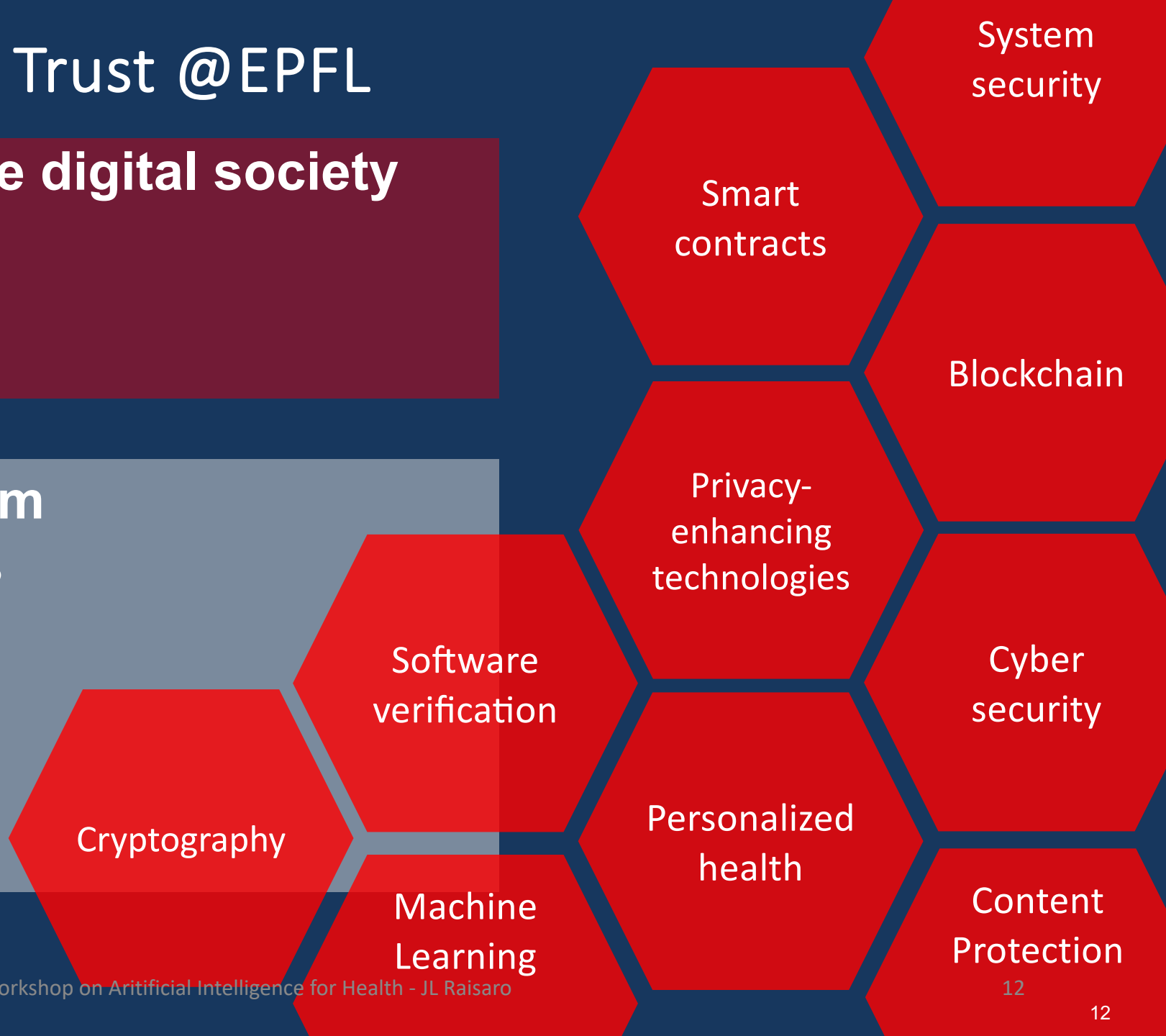
The Center for Digital Trust @EPFL

Reinventing trust for the digital society

- Center of expertise
- One-stop-shop
- Community

Collaboration Ecosystem

- 30+ EPFL laboratories
- 8+ Organizations



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

https://c4dt.org/

Thank you!