Dr. Roberto Yus
https://robertoyus.com/

Data Privacy

CMSC 463/663

L03 – Privacy Enhancing Technologies

Time to adopt PETs!*

* https://www.enisa.europa.eu/topics/data-protection/privacy-enhancing-technologies

Previously on…

- Privacy by design
- Proactive, by default, embedded, end-to-end, user centric
- We need PETs to implement it!

‘A privacy nightmare’: the $400m surveillance package inside the US immigration bill

Experts issue warning over bipartisan measure’s funding for towers and DNA tests that would ‘hyper-amplify what’s already happening’

In the news!

Privacy Enhancing Technologies

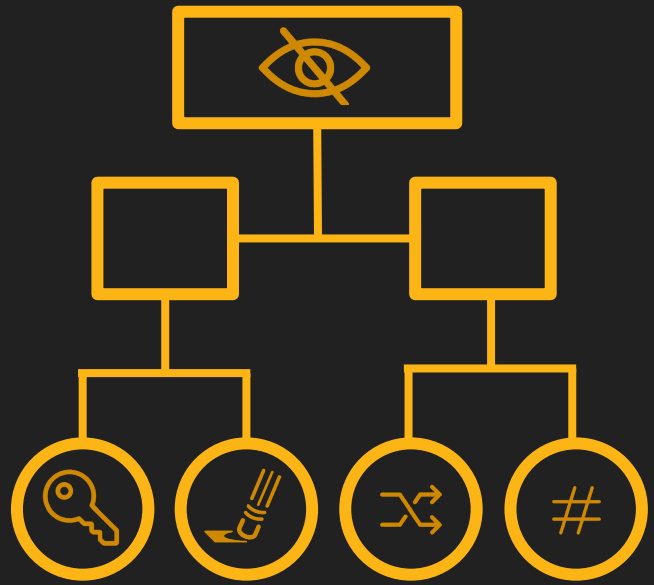
*"coherent system of **ICT measures that protects privacy** by **eliminating or reducing personal data** or by **preventing unnecessary and/or undesired processing of personal data**; all without losing the functionality of the data system"* [1]

*"wide range of **technologies that help protect personal privacy**. Ranging from tools that provide **anonymity** to those that allow a user to choose if, **when** and under what circumstances personal information is **disclosed**, the use of privacy enhancing technologies helps users make **informed choices about privacy protection**."* [2]

[1] John J. Borking and Charles D. Raab. Laws, PETs and other Technologies for Privacy Protection. Journal of Information, Law & Technology (JILT), 1(1), 2001.
[2] Organisation for Economic Co-operation and Development (OECD). Inventory of Privacy-Enhancing Technologies (PETs). Report DSTI/ICCP/REG(2001)1/FINAL, working party on information security and privacy. Technical report, 2002.

PETs Classification

- **Classification is hard!**
  - Soft PETs vs. Hard PETs
  - Anonymity vs. Unlinkability vs. Undetectability vs. Unobservability vs. Pseudonymity vs. Identity Management
  - Cryptographic vs. Masking vs. Others
  - …



Examples of PETs

- **Traditional PETs**
  - Encryption
  - De-Identification
  - Access Control
  - …
- **Emerging PETs**
  - Homomorphic encryption
  - Trusted execution environment
  - Differential privacy
  - Multi-party computation
  - Federated analytics
  - …

Examples of PETs

- **Traditional PETs**
  - Encryption
  - *De-Identification* **L06!**
  - *Access Control* **L05!**
  - …
- **Emerging PETs**
  - Homomorphic encryption
  - Trusted execution environment
  - *Differential privacy* **L07!**
  - *Multi-party computation* **L08!**
  - *Federated analytics* **L09!**
  - …

Encryption

- One of the **principal security technologies** to protect information
- Converts legible data into **ciphertext**
  - **Unreadable by a human or a computer**
- Data can only be **read by first decrypting it**
  - Requires access to **decryption key**
- Types of encryption:
  - **In transit** → secures data as it flows between two connected computers.
  - **At rest** → secures data for storage on disk

Encryption

Limitations?

- Pros: Mature technologies, widely used
- Cons: Whoever has access to the key can see all data,…

De-Identification Techniques

- **Reduces amount of information** about individual in a dataset, and/or **reduces risk of re-identification** of individual
- Involve direct **manipulation of raw data**
- Examples:
  - **Redaction**: deleting an entire record or field, or obfuscating part of a record or field
  - **Tokenization**: replacing a real value with a randomly generated value
  - **Hashing**: applying a function to a value to produce a fixed-length value (or hash)
  - **Generalization**: transforming a value to a less precise or bucketed value
  - **k-anonymity**: any record in the dataset becomes indistinguishable from (k-1) records

https://cdeiuk.github.io/pets-adoption-guide/

De-Identification Techniques

Limitations?

- Pros: Data can be shared
- Cons: No matter what technique is used there is risk of re-identification!...

De-Identification Techniques

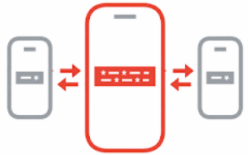
Google

Privacy-Preserving Contact Tracing

Across the world, governments and health authorities are working together to find solutions to the COVID-19 pandemic, to protect people and get society back up and running. Software developers are contributing by crafting technical tools to help combat the virus and save lives. In this spirit of collaboration, Google and Apple are announcing a joint effort to enable the use of Bluetooth technology to help governments and health agencies reduce the spread of the virus, with user privacy and security central to the design.

Once you opt-in to the notification system, the Exposure Notifications System will generate a random ID for your device. To help ensure these random keys can't be used to identify you or your location, they change every 10-20 minutes.

Your phone and the phones around you will work in the background to exchange these privacy-preserving random keys via Bluetooth. You do not need to have the app open for this process to take place.



https://covid19.apple.com/contacttracing
https://www.google.com/covid19/exposurenotifications/#exposure-notifications-and-privacy

Homomorphic Encryption (HE)

- **Computation directly on encrypted data**
  - Traditional encryption → in transit & at rest
  - HE → **in process**
- HE schemes:
  - **Partial homomorphic encryption (PHE):** permits only a single type of operation (e.g., addition) on encrypted data.
  - **Somewhat homomorphic encryption (SHE):** permits some combinations of operations (e.g., some additions and multiplications) on encrypted data.
  - **Fully homomorphic encryption (FHE):** permits arbitrary operations on encryption data.

https://cdeiuk.github.io/pets-adoption-guide/

Homomorphic Encryption (HE)

Limitations?

- Pros: Nobody else (apart from me) can decrypt the data
- Cons: Significant computation overhead (FHE), support for limited operations, debugging is complex,…

Homomorphic Encryption (HE)

ENVEIL
ENCRYPTED VEIL

Enveil and DeliverFund Leverage Privacy Enhancing Technologies to Combat Human Trafficking

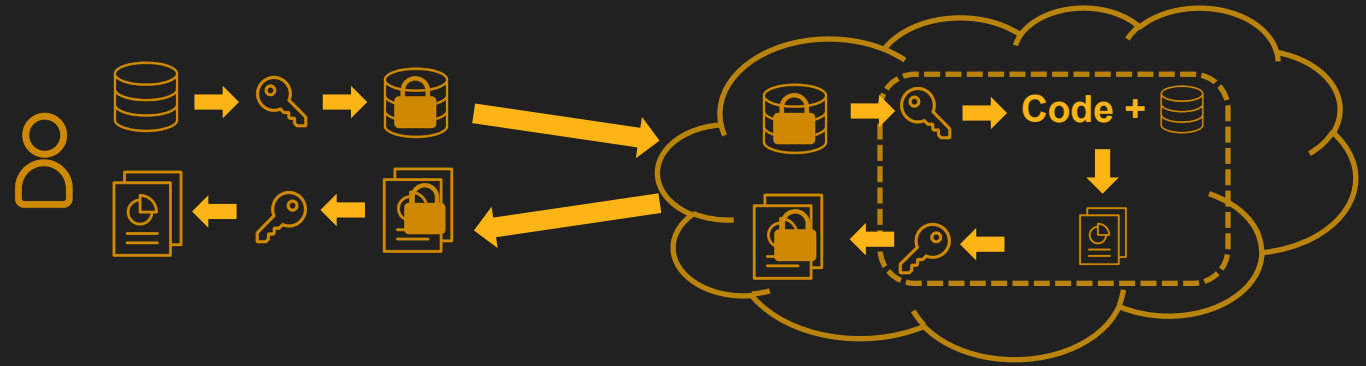
Partnership demonstrates the power of utilizing data and technology as a force for social good

June 15, 2021 09:00 ET | Source: Enveil

processing. Organizations can submit sensitive queries through PATHFinder Advantage with confidence knowing that sensitive customer data will never be exposed in clear text. This enhanced security posture, made possible by Enveil's advances in homomorphic encryption, allows entities to share and collaborate with data in a secure and private capacity that was never before possible.

https://www.globenewswire.com/en/news-release/2021/06/15/2247363/0/en/Enveil-and-DeliverFund-Leverage-Privacy-Enhancing-Technologies-to-Combat-Human-Trafficking.html

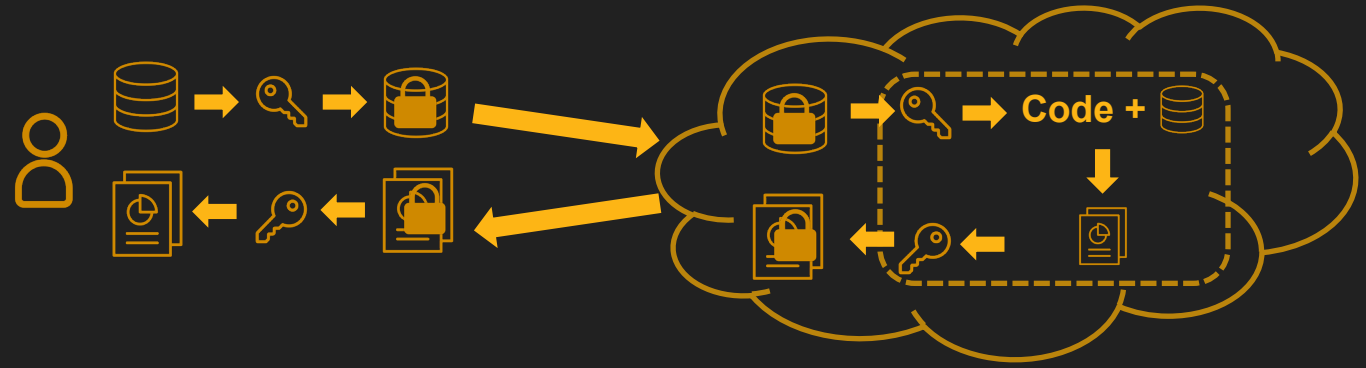
Trusted Execution Environment (TEE)



- TEE → processing environment **isolated** from a computer's main processor and memory
- Code/data **cannot be accessed** by main processor
- **Encrypted communications** between main processor and TEE

https://cdeiuk.github.io/pets-adoption-guide/

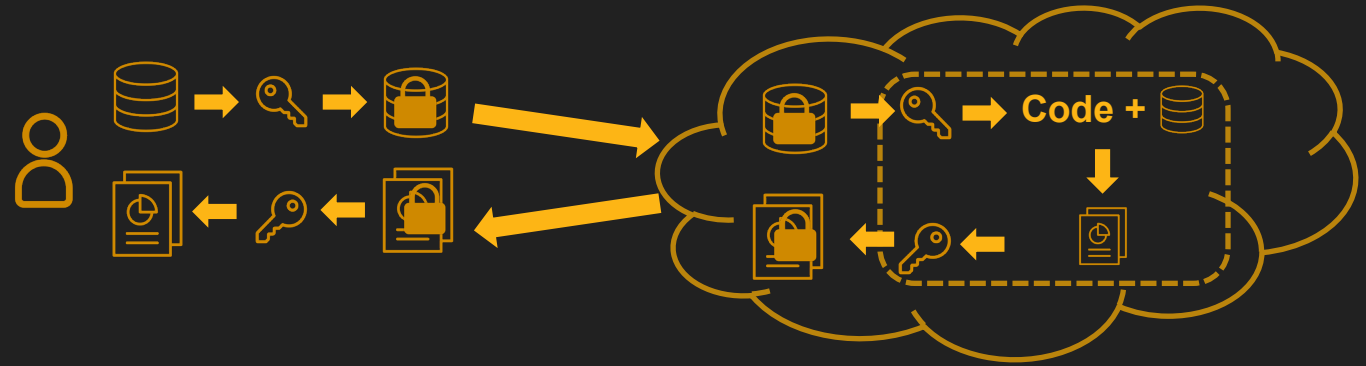
Trusted Execution Environment (TEE)



Limitations?

- Pros: Ability to perform more complex operations in trusted domain
- Cons: Potential side-channel attacks, hardware limitations,…

Trusted Execution Environment (TEE)



Technology preview: Private contact discovery for Signal

moxie0 on 26 Sep 2017

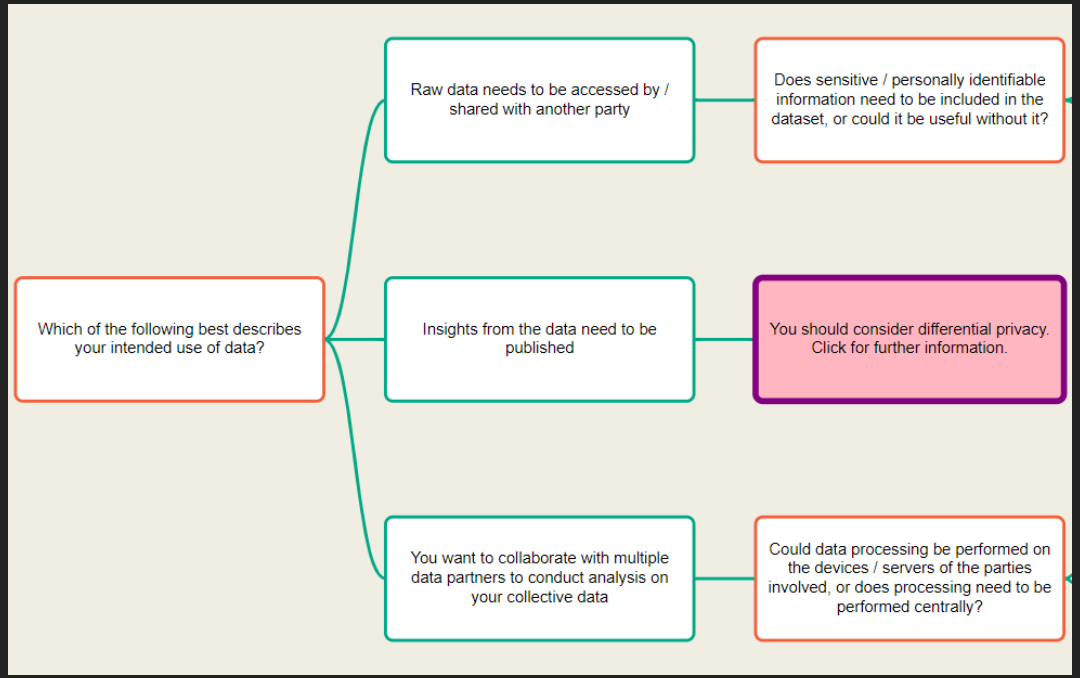
Signal

Modern Intel chips support a feature called Software Guard Extensions (SGX). SGX allows applications to provision a "secure enclave" that is isolated from the host operating system and kernel, similar to technologies like ARM's TrustZone. SGX enclaves also support a feature called *remote attestation*. Remote attestation provides a cryptographic guarantee of the code that is running in a remote enclave over a network.

https://signal.org/blog/private-contact-discovery/

Choosing PETs

- **No silver bullet!**
- Think about different data dimensions:
  - What data are you handling?
  - What's the intended use?
  - Accessed by other parties?
  - Publicly published?
  - Can data be processed locally?
  - …



https://cdeiuk.github.io/pets-adoption-guide/

Discussion

Case Study: Smart Campus

- Greendale Community College (GCC) wants to implement a new system:
  - Students pay only for the classes they attend
  - Faculty salary depends on how many students attend their classes
  - ....

What extra functionalities can you offer?

What PETs would you use?

https://cdeiuk.github.io/pets-adoption-guide/adoption-guide