**Privacy Enhancing Technologies**

Privacy-enhancing technologies (PETs) are a broad range of technologies (hardware or software solutions) that are designed to extract data value in order to unleash its full commercial, scientific and social potential, without risking the privacy and security of this information.

# **Importance of PETs**

Like any other data privacy solution, privacy-enhancing technologies are important due to three reasons for businesses:

* Data protection laws such as GDPR and CCPA are forcing organizations to preserve consumer data. Businesses can pay serious fines because of data breaches.
* Data may need to be tested by third-party organizations due to the lack of your business’ self-sufficiency in analytics and application testing. PETs enable privacy protection while data sharing.
* Privacy breaches can harm your business’ reputation, businesses or customers (depending on your business model) may want to stop interacting with your brand.An example is the share price loss of Facebook after Cambridge Analytica scandal.

# **Examples of common privacy-enhancing technology**

## **Cryptographic algorithms**

* Homomorphic Encryption: Homomorphic encryption is an encryption method that enables computational operations on encrypted data. It generates an encrypted result which, when decrypted, matches the result of the operations as if they had been performed on unencrypted data (i.e. plaintext). This enables encrypted data to be transfered, analyzed and returned to the data owner who can decrypt the information and view the results on the original data. Therefore, companies can share sensitive data with third parties for analysis purposes. It is also useful in applications that hold encrypted data in cloud storage. Some common types of homomorphic encryption are:
  + Partial homomorphic encryption: can perform one type of operation on encrypted data, such as only additions or only multiplications but not both.
  + Somewhat homomorphic encryption: can perform more than one type of operation (e.g. addition, multiplication) but enables a limited number of operations.
  + Fully homomorphic encryption: can perform more than one type of operation and there is no restriction on the number of operations performed.
* Secure multi-party computation (SMPC): This is a subfield of homomorphic encryption with one difference: users are able to compute values from multiple encrypted data sources. Therefore, machine learning models can be applied to encrypted data since SMPC is used for a larger volume of data.
* Differential privacy: Differential privacy protects from sharing any information about individuals. This cryptographic algorithm adds a “statistical noise” layer to the dataset which enables to describe patterns of groups within the dataset while maintaining the privacy of individuals.
* Zero-knowledge proofs (ZKP): ZKP uses a set of cryptographic algorithms that allow information to be validated without revealing data that proves it.

## **Data masking techniques**

Some privacy enhancing technologies are also data masking techniques that are used by businesses to protect sensitive information in their data sets.

* Obfuscation: This one is a general term for data masking that contains multiple methods to replace sensitive information by adding distracting or misleading data to a log or profile.
* Pseudonymization: Identifier fields (fields that contain information specific to an individual) are replaced with fictitious data such as characters or other data. Pseudonymization is frequently used by businesses to comply with GDPR.
* Data minimisation: Collecting a minimum amount of personal data that enables the business to provide the elements of a service.
* Communication anonymizers: Anonymizers replace online identity (IP address, email address) with disposal/one-time untraceable identity.

## **With the help of AI & ML algorithms**

* Synthetic data generation: Synthetic data is an artificially created data by using different algorithms including ML algorithms. If you are interested in privacy-enhancing technologies because you need to transform your data into a testing environment where third-party users have access, generating synthetic data that has the same statistical characteristics is a better option.
* Federated learning: This is a machine learning technique that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging them. With the decentralization of servers, users can also achieve data minimization by reducing the amount of data that must be retained on a centralized server or in cloud storage.

# Uses of PETs

* Test data management: Application testing and data analysis are sometimes handled by third-party providers. Even when they are handled in-house, companies should minimize internal access to customer data. Using a suitable PET that doesn’t significantly affect test results is important for organizations.
* Financial transactions: Financial institutions are responsible for protecting the privacy of the customers due to citizens’ freedom to conduct private deals and transactions with other parties.
* Healthcare services: Healthcare industry collects and shares (when needed) electronic health records (EHR) of patients. For example, clinical data can be used for searching for adverse effects of various drug combinations. Healthcare companies ensure the privacy of patients’ data in such cases by using PETs.
* Facilitating data transfer between multiple parties including intermediaries: For businesses that work as a middle man between two parties, the usage of PETs is crucial since these businesses are responsible for protecting the privacy of both parties’ information.