MASON HAYES & CURRAN

# Privacy Enhancing Technologies - Key Considerations

10 October 2024

The Information Commissioner (ICO) recently published its report[1] on 'privacy-enhancing technologies' (PETs). In this report the ICO analyses the challenges hindering the adoption of PETs and provides recommendations to foster responsible PETs across various sectors.

The ICO also published detailed guidance on PETs last year[2]. Given PETs are aimed at assisting controllers achieve compliance with the GDPR and other privacy focused legislation, these suite of tools are pivotal for use in data-driven technologies such as AI.

We highlight the key learnings from these regulatory sources, enabling controllers to avail of AI's capabilities whilst also achieving compliance with the GDPR.

## What are PETs?

PETs embody fundamental data protection principles by:

- Minimising personal information use
- Maximising information security, and
- Empowering people[3]

The term is not defined within the GDPR. However, the European Union Agency for Cybersecurity (ENISA) has defined the concept as:

> *"Software and hardware solutions, i.e. systems encompassing technical processes, methods, or knowledge to achieve specific privacy or data protection functionality or to protect against risks of privacy of an individual or a group of natural persons".*

PETs can help reduce the risks to rights and freedoms that a processing activity may pose. For example, these tools can be a suitable technical and organisational measure for the types of processing likely to result in a high risk. In particular, for processing that involves large-scale data collection, e.g. AI, Internet of Things, and cloud computing services.

Taking AI processing as an example, PETs can aid compliance by:

| Processing activity | Risks | PETs which may aid compliance |
| --- | --- | --- |
| Processing involving artificial intelligence, machine learning, and deep learning applications. | Possible risks to people involved in the training dataset include model inference and attribute inference attacks.<br><br>These can reveal individuals' identities or may result in learning sensitive information about them. | PETs can help assess and mitigate these risks.<br><br>For example:<br><br>- Encryption ensures that only parties with the decryption key can access the information. This protects the information that is being processed, e.g. to train the AI model.<br><br>- Secure multi-party computation (SMPC) is a protocol that distributes a computation across multiple parties and where no individual party can access the other parties' data. As a result, data stored in shared global models can be individually protected.<br><br>- Differential privacy adds random noise during training to ensure the final model does not memorise information unique to a particular individual's personal information.<br><br>- Federated learning can minimise the amount of centrally held personal information and reduce the transfer of personal information between parties, and<br><br>- Synthetic data can be used at the training stage to reduce the amount of personal information used to train AI[4]. |

While these tools are not new, the recent increase of machine learning systems have brought them into focus. The rise of the data economy and increased privacy awareness among end-users, have also contributed to their growing importance.

## The ICO's findings

The ICO's multi-stakeholder review in the adoption of PETs resulted in the below findings[5]:

| Barrier to adoption | Industry comments |
|---|---|
| Limited awareness of PETs | Participants emphasised a real need for industry to gain a greater understanding of PETs, how they are defined, and where they fit into broader data governance frameworks. |
| Lack of technical expertise | Insufficient understanding from PET providers of the environments in which their products will be deployed within. |
| Uncertainty on the maturity of PETs | Data sharers agreed that the PETs market needed to mature further to enable widespread adoption. |
| Unclear and complex pricing. As well as a lack of understanding of the costs and benefits analysis relating to PETs | Uncertainty regarding the costs of PETs products was seen as a significant issue, both in terms of upfront procurement and ongoing costs associated with maintenance of the service. |
| Regulatory uncertainty around the use of PETs for compliance | Challenges around assessing identifiability of data when using PETs was a common issue. Developers focused on the ability of their PETs products to anonymise or reduce identifiability, in order to reduce or remove compliance.<br><br>However, data sharers often required re-identification and broadly described PETs as an enhancement to their data governance structure, as opposed to a means to avoid compliance responsibilities. |

## Understanding PETs

One of the biggest challenges in adopting PETs, as found in the ICO's report, was a lack of understanding. Referring back to the ICO's detailed guidance on PETs, we briefly discuss what these tools are and how they can assist with GDPR compliance.

PETs are available for a variety of purposes, for example:

- Secure training of AI models
- Generating anonymous statistics, and
- Sharing information between different parties

Examples of PETs include[6]:

| Altering Data | Shielding Data | Systems + Architecture |
|---|---|---|
| Anonymisation | Encryption | Multi-party Computation |
| Pseudonymisation | Homomorphic Encryption | Data Dispersion |
| Differential Privacy | Privacy Enhanced Hardware | Management Interfaces |
| Synthetic Data | | Digital Identity |

In the context of AI and other data-driven technologies, using privacy preserving techniques such as anonymisation, differential privacy and synthetic data can be particularly useful in achieving compliance with the GDPR.

## Anonymisation

Fully anonymised data renders an individual unidentifiable. As a result, the data does not fall under scope of the GDPR. Therefore, anonymisation is not a PET. This is because, the purpose of many PETs is to enhance privacy and protect personal data processed, rather than anonymise it entirely.

That said, PETs and anonymisation are related concepts. While not all PETs result in effective anonymisation, some can play a role in anonymisation. For example, differential privacy methods can enable controllers to prevent information about specific individuals from being revealed or inferences about them being made.

## Differential privacy

Differential privacy is a property dataset or database, providing a formal mathematical guarantee about people's indistinguishability. These solutions inject noise into a dataset or into the output of a machine learning model, without introducing significant negative effects on data analysis or model performance. The end result is a differentially private dataset or model that cannot be reverse engineered by an attacker. As a result, adoption of this tool can make it impossible to identify with certainty individual records e.g. customers, patients, within a dataset. For example, analysing driving information for every driver in Dublin without being able to identify any individual driver. In this example, macro level questions about driver behaviour and road safety can be answered without compromising the individual privacy of the individuals contributing to the data.

Differential privacy can:

- Anonymise personal data, once enough noise is added, and/or

- Query a database to provide anonymous information (statistics)

Differential privacy does not necessarily result in anonymous information. If not properly configured, there is a risk of personal data leakage from a series of different queries. For example, if the privacy budget is poorly configured, an attacker can accumulate knowledge from multiple queries and re-identify someone.

## Synthetic data

Synthetic data is 'artificial' data generated by data synthesis algorithms. It replicates patterns and the statistical properties of real data, which may be personal data. It is generated from real data using a model trained to reproduce its characteristics and structure. This means that outputs using synthetic data should be largely the same as those derived from using the original real data[7].

This PET is particularly useful for training AI models. This PET enables controllers to generate large datasets from small datasets. As a result, this PET enables controllers to comply with data minimisation as it reduces processing of personal data.

Generating synthetic data from personal data will carry through any inherent biases from the original dataset. As a result, it is important that biases are detected and corrected in the generation of synthetic data. To mitigate against these risks, a diverse and representative training dataset should be used when generating synthetic data.

## Recommendations of the ICO

The ICO held an in-person workshop in February of this year, to gain a deeper understanding of the barriers to adoption for both suppliers and users of PETs. We summarise what the ICO intends to do to overcome these challenges:

1. Integrate its guidance on PETs into other guidance, e.g. to reduce risks to people in AI use cases.
2. Publish further PET case studies to demonstrate best practice for potential adopters.
3. Contribute to the development of standards and accreditations for PETs with relevant stakeholders.
4. Clarify when PETs can provide effective anonymisation and pseudonymisation.
5. Additional guidance before year end is expected and which will detail the cost benefit analysis for various PETs.

## Conclusion

The key takeaways from the ICO's report make clear that there is a significant lack of awareness around PETs. This should be addressed, particularly as PETs can be helpful in demonstrating a 'data protection by design and default' approach has been adopted, specifically with regards to AI related processing.

That said, there are risks associated in utilising these solutions - particularly if not correctly deployed - given how technical they are by nature. As a result, controllers should obtain legal and technical advice as appropriate .

For more information and expert advice on PETs, please contact a member of our Privacy & Data Security team.

# People also ask

**What are PETs?**

PETs such as end-to-end encryption, are a broad set of tools and methods aimed at providing ways to build products and functionality while protecting the privacy of users' data.

**How do PETs relate to data protection law?**

PETs are linked to the concept of 'data protection by design and default' and are therefore relevant to the technical and organisational measures adopted. Adoption of PETs can help controllers achieve compliance with the data protection principles, and are viewed as mitigating measures, decreasing the risk of processing personal data.

**What are the risks of using PETs?**

PETs are not a "silver bullet" to compliance with the GDPR. Any processing which has utilised PETs into the design, must still be lawful, fair and transparent. Before considering PETs, controllers should:

- Assess the impact of the processing
- Understand the processing purpose
- Understand and document how PETs can assist in achieving compliance with the GDPR, and
- Consider and address the issues PETs may pose to compliance with other data protection principles, e.g. issues with accuracy and accountability.

*The content of this article is provided for information purposes only and does not constitute legal or other advice.*

---

[1] On X (formerly known as Twitter), the ICO announced that is had published a report highlighting its key findings and recommendations following its ongoing PETs work and a workshop it hosted to gain a deeper understanding of the barriers that exist for users and developers of PETs (the ICO's report), 13 August 2024, available here.

[2] The ICO, '*Privacy-enhancing technologies (PETs)*' (the ICO's guidance), June 2023.

[3] The ICO's guidance, page 4.

[4] The ICO's guidance, 'Which types of processing can benefit from using PETs?, page 11.

[5] The ICO's report.

[6] Image taken from: Federal Reserve Bank of San Francisco, 'Privacy Enhancing Technologies: What Are They and Why Do They Matter', 3 June 2021, available here.

[7] The ICO's guidance, page 27.