Blueprints for Deploying Privacy Enhancing Technologies in E-Government

Liina Kamm(✉), Dan Bogdanov, Eduardo Brito, and Andre Ostrak

Cybernetica AS, Mäealuse 2/1, 12618 Tallinn, Estonia
liina.kamm@cyber.ee

Abstract. Governments are increasingly becoming providers of data-driven services to both citizens and organisations. As the number of these services grows, the government will store greater amounts of personal and company data. Data minimisation and data protection can be legal obligations, especially for governments that have passed data protection and privacy regulations. However, even without such regulation, processing only as much data as is necessary, is good information security practice. In this paper, we give an overview of how we put together a Privacy Enhancing Technology (PET) concept and roadmap for an e-government, including motivations, an adoption strategy and blueprints for services that benefit from PETs.

Keywords: E-government · services · privacy enhancing technologies

1 Introduction

Privacy enhancing technologies (PETs) are privacy controls that reduce personally identifiable information (PII) or prevent unnecessary processing of PII without losing the functionality of the system. PETs can be categorised according to the privacy goals that they achieve—transparency, intervenability and unlinkability. Transparency means the availability of comprehensive information about data processing. Intervenability means that the data subject has the opportunity to control data processing. A vast majority of contemporary, deployable PETs target unlinkability, striving to prevent the re-identification of data subjects.

Accomplishing these privacy goals requires aligning the organisation's operations with privacy principles. The goals must be considered in the early stages of system development. Choosing the right PET for the task is a key challenge. Moreover, some technologies, e.g., logging, are not directly PETs but their correct use allows the system developer to achieve privacy goals.

The development of PETs has not been a priority when compared to the rate of development of IT systems in general. The adoption rate of PETs is low. Hence, stronger incentives are needed to increase the uptake. Data protection regulations in different territories approach privacy in different ways. None directly require PET deployments, but instead compel data processors to deploy

F. Bieker et al. (Eds.): Privacy and Identity 2023, IFIP AICT 695, pp. 3–19, 2024.
https://doi.org/10.1007/978-3-031-57978-3_1

controls on par with the state of the art. Thus, it is important for PETs to become the state of the art.

This paper reports on the authors' experience of putting together the PET strategy Estonia. Estonia is a mature e-government according to the United Nations e-Government Survey [6]. In 2022, the Ministry of Economic Affairs and Communications initiated a project to prepare the PET strategy for e-government. In 2023, the concept [4] and roadmap [5] reports were published. In this paper, we focus on the benefits of using PETs (Sect. 2) and give a survey of international use of PETs (Sect. 3). In Sect. 4 we give recommendations on how to include PETs in the development lifecycle and, finally, in Sect. 5, we propose PET-powered service blueprints for an e-government.

2 Expected Benefits of PETs in Digital Societies

PETs are a Tool for Empowering the Participants of a Digital Society. An emerging open digital society is a distinct value, and within it, individuals can become advocates for innovation—if given the tools for maintaining their data autonomy. Individuals should be empowered to actively participate in ensuring their own privacy. If citizens, companies, and the government can engage with each others' data processing applications without being afraid that the data they grant will be used against them, it will encourage the uptake of new services. For this to be possible, a sufficient maturity among both service providers and users is needed.

PETs Facilitate the Creation of New Services. There is potential in the reuse of public sector data. Based on their data, individuals can be offered event-based services, needs-based subsidies, or tailored services. Data-driven policy-making and the automation of processes in such services may result in a more efficient use of resources in the public sector. For the private sector, public sector data can serve as a reliable source upon which to build new services.

PETs could greatly benefit healthcare, finance, education, internet marketing, the Internet of Things, and public administration. For example, health data are a special data category requiring stronger protection. Hence, data use for proactive services is limited. However, personal health data have value for medical research, public health policy-making, and e-health application development. With PETs, the reuse of health data does not have to compromise privacy.

The principle of data protection by design and by default implies that the appropriate technical measures must be applied, especially in high-risk data processing scenarios. If properly implemented, PETs are such measures. Raising awareness about PETs and sharing successful application stories and best practices is of critical importance. Ethical and legal questions need to be considered together with PET use.

PETs Support Data Protection by Design and by Default. Article 25 of the European General Data Protection Regulation (GDPR) [17] mandates that data controllers implement data protection by design and by default when processing personal data. These are complementary concepts that enhance data protection. Using PETs is one way of achieving this requirement, but other techniques, like privacy engineering, are also required. We will discuss this in Sect. 4.

PETs Protect Both Personal and Organisational Data. Business or organisational data, including trade secrets that do not contain personal data, are not subject to the same privacy requirements as personal data. PETs can also provide enhanced protection for other types of confidential data (e.g., trade secrets), thereby reducing risks for organisations and positively impacting the development of services for, e.g., real-time economy, taxation, fraud detection.

Companies using PETs may gain greater trustworthiness in the eyes of clients and consumers. Additionally, PETs can help companies mitigate risks and reduce the number of data breaches and unauthorised accesses. Conversely, neglecting these measures can result in undesirable risks and real economic losses.

3 Existing Deployments of Privacy Enhancing Technologies in E-Government

In this section, we give an overview real-life deployments of PETs in governments. We focus on the deployment of complex PETs in e-government services and strategies, and leave out simpler technologies such as pseudonymisation.

3.1 Canada

PETs in Government Strategy. The concept of privacy-by-design originated in the Office of the Privacy Commissioner of Canada. In privacy-by-design systems, privacy requirements are considered throughout the system's entire lifecycle. They have also released a comprehensive PETs review [22].

PETs in E-Government Services and Regulation. Statistics Canada has developed a proof of concept for private set intersection combined with secure multiparty computation, and a proof of concept for training machine learning models on synthetic data. In 2018 and 2019, Statistics Canada generated synthetic datasets from real data using a mass data imputation method. The data was used in hackathons and other training exercises [26].

3.2 Estonia

PETs in Government Strategy. The Estonian Digital Agenda 2030 [14] highlights the protection of fundamental human rights, including privacy, as a primary principle. The plan envisions the development of a human-centric digital state as the next significant leap in digital society. Concrete actions include raising awareness, and a more wide-spread implementation of data trackers and consent services. The plan emphasises the significance of implementing PETs to enable data reuse and data-driven governance. A national program for the application of

PETs is set to be launched. Through these technologies, Estonia aims to achieve higher levels of transparency, intervenability, and unlinkability.

PETs in E-Government Services and Regulation. The Estonian internet voting system IVXV employs mix networks and zero-knowledge during the vote-tallying phase to de-identify signed e-votes [11]. Estonia has deployed secure multi-party computation to enable data analysis when anonymisation could not be used [2]. The Estonian government is offering an e-government-wide transparency service that allows citizens to see which organisations have been accessing their records. There is also a public catalogue for e-government information systems and data.

3.3 France

PETs in Government Strategy. The French data protection agency CNIL has published materials on PETs, e.g., pseudonymisation and anonymisation techniques [1]. CNIL regularly organises a scientific conference called the Privacy Research Day, which features PETs. CNIL has also published a methodology and guidelines for privacy impact assessment [3].

3.4 Japan

PET Research and Standardisation. The Japanese Institute of Legal Informatics and Systems (JILIS) focuses on research and policies related to information systems. The institute publishes recommendations on data protection, highlighting deficiencies or ambiguities in documents related to data storage and sharing, including system specifications, reports, laws and bills. They often emphasise the need to distinctly differentiate between anonymisation, pseudonymisation and encrypted data, addressing the limitations and opportunities of each technique.

Industry PET Initiatives. In August 2022, Japan established a voluntary consortium known as the *Privacy Tech Association*[1]. The primary objective of this association is to promote the use of privacy technologies in companies working with data. Research areas include anonymisation, secure multi-party computation, differential privacy, and data synthesis.

3.5 Netherlands

PETs in Government Strategy. Under the leadership of the Netherlands Ministry of Economic Affairs and Climate Policy, a cryptographic roadmap is being developed, which also encompasses PETs [13]. The document highlights opportunities for using these technologies in the healthcare sector, facilitating better and privacy-preserving use of sensitive data for disease research.

PETs in E-Government Services and Regulation. The Netherlands' statistics office CBS has been studying PETs since 2018 to support secure and sustainable

[1] https://privacytech-assoc.org (last checked 10.10.2023).

data sharing [19]. In collaboration with the health insurance company CZ and Zyderland Hospital, CBS has also created a pilot platform that uses private set intersection, homomorphic encryption, and secure multi-party computation to link and analyse health data [27].

3.6 Singapore

PETs in Government Strategy. In 2022, the government agency IMDA launched a secure testing sandbox for pilot projects related to PETs [9]. The aim is to identify technologies that can assist companies with data sharing challenges and to understand their limitations. Based on the outcomes of the pilot projects, IMDA and the data protection agency PDPC plan to identify feasible PETs and develop standards and policies for the adoption of these technologies.

PETs in e-Government Services and Regulation. During the COVID-19 pandemic, Singapore pioneered TraceTogether, an app utilising Bluetooth Low Energy technology for contact tracing [20]. The solution was praised for not relying on privacy-infringing techniques like GPS tracking.

3.7 Switzerland

PETs in E-Government Services and Regulation. Switzerland has been developing their internet voting technology. In 2021, the source code and documentation were released to allow volunteers to test the system. The new system employs end-to-end encryption, mix networks, and zero-knowledge proofs for preserving voters' privacy. At the start of the voting process, the vote is encrypted and sent to the voting server. Mix networks are used to shuffle the votes before tallying, the process is overseen by independent auditing components. Zero-knowledge proofs ensure that during the mixing and decryption processes, votes are not added, deleted, or altered [21]. Researchers from the Swiss Federal Institutes of Technology, EPFL and ETH Zurich, helped develop the DP-3T technology for contact tracing of COVID-19 patients. DP-3T gained widespread adoption due to its openness, privacy guarantees and decentralised approach [25]. The Swiss Academy of Medical Sciences (SAMS) oversees the Swiss Personalized Health Network that allows analysts to conduct research across data from multiple institutions by using homomorphic cryptography, secure multi-party computation, and differential privacy [16].

3.8 United Kingdom

PETs in Government Strategy. In 2022, the United Kingdom and the United States started a PET prize challenge [24] to encourage technology uptake. The Centre for Data Ethics and Innovation (CDEI) spearheaded the organisation and planning of the challenge from the UK side.

PETs in E-Government Services and Regulation. The Office for National Statistics (ONS) has been studying synthetic data generation since 2018, with the aim of testing machine learning and data pipelines. In preparation for the 2021 census, the ONS generated synthetic data for testing data analysis and balancing workloads. ONS is also exploring the use of differential privacy in data synthesis [26].

PET Research and Standardisation. In 2023, the British Royal Society published a report on PETs [23]. They explored the public sector's interest in PETs and found that the surveyed institutions were mainly interested in pseudonymisation, anonymisation, data synthesis, differential privacy, and federated learning.

3.9 United States of America

PETs in Government Strategy. The White House Office of Science and Technology Policy, the National Institute of Science and Technology and the National Science Foundation led the planning on the US side [24].

PETs in E-Government Services and Regulation. A study of US Census has shown that anonymisation does not prevent re-identification [18]. The paper shows that 99.98% of Americans can be re-identified based on 15 demographic features. The US Census Bureau itself was able to re-identify the geographic location, gender, age, racial, and ethnic affiliation of 52 million anonymised individuals [7]. This led the bureau to adopt differential privacy for the 2020 census. The US Census Bureau has previously also used synthetic data generation to replace unique values with generated ones [12]. Higher education is often the most significant investment in the lives of US citizens, in addition to purchasing a home. To help prospective students make better decisions when pursuing higher education, the Student Right to Know Before You Go Act of 2019 was presented to the US Congress [8]. The aim of the legislation is to establish a more accurate, comprehensive, and privacy-preserving data system that allows inquiries about higher education institutions.

PET Research and Standardisation. The onion routing technology was developed in the US Naval Research Laboratory at the end of the 20th century. Subsequently, the technology was further developed by the Defense Advanced Research Projects Agency (DARPA). US research agencies such as DARPA, IARPA, and NSF systematically fund programs that develop PETs (e.g., fully homomorphic encryption, secure multi-party computation, zero-knowledge proofs) with applications in various domains. The national standardisation agency NIST is working on several privacy initiatives, including technology reports.

4 Privacy in E-Government Service Development

4.1 Applying Privacy Engineering in the Service Lifecycle

Privacy engineering is a discipline aimed at reducing privacy risks, justifying the allocation of developmental resources, and implementing protective measures in information systems. Privacy engineering supports compliance with privacy goals

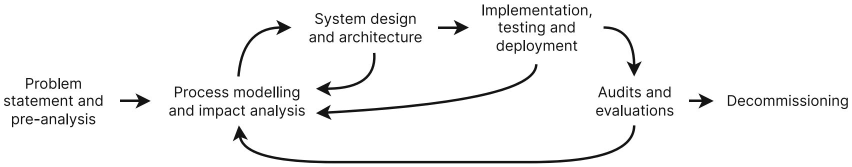
Problem statement and pre-analysis
Process modelling and impact analysis
System design and architecture
Implementation, testing and deployment
Audits and evaluations
Decommissioning

Fig. 1. Simplified lifecycle of an e-government service

and data protection requirements. Privacy engineering can include requirements analysis, functional modelling, risk and impact assessment, architectural design and system implementation.

The same stages are present in the lifecycle of a system (e.g., an e-government service). Each system exists at some stage within this lifecycle. For instance, an operational information system may be assessed at a given point to decide whether to continue its development or decommission it. At each of these steps, decisions must be made and steps must be taken to better achieve the system's privacy goals. Figure 1 shows a simplified lifecycle of an e-government service.

The cyclical nature implies that there must be a way to revisit earlier stages. We propose that the development and the operation stage can revert back to the analysis stage, if an internal requirements (e.g., emergence of new requirements during development) or external requirements (e.g., audits, legal assessments, opportunities for system improvement) demands.

The ISO/IEC TR 27550 [10] technical report aligns privacy engineering, system development, security engineering and risk management. The approach is generic and aligns with other lifecycle management practices and standards.

4.2 Integrating PETs in Service Development

Problem Statement and Pre-analysis. Relevant activities include:

- enumerating stakeholders and data flows at a high level, and
- formulating requirements for minimising data processing.

Specific PETs do not need to be selected at this stage. Instead, requirements for further development should be formulated. This stage may be divided into substages, carried out by different stakeholders. For instance, a governmental body may conceptualise the task, but a service provider may perform the preliminary analysis through a procurement (see Sect. 4.3).

Process Modelling and Impact Analysis. Relevant activities include:

- refining data flow and stakeholder specifications,
- collecting data protection requirements,
- identifying stakeholders' rights and capabilities for personal data processing,
- conducting an initial data protection impact assessment (DPIA).

The goal is to create models with sufficient precision to identify the data elements each stakeholder processes, and to determine the level of identifiability required for fulfilling the system's objective. Assigning data controller and processor responsibilities and understanding the stakeholders' technical capabilities in deploying data processing technologies helps design the system. Based on this information, an initial data protection impact assessment can be performed.

System Design and Architecture. Relevant activities for design include:

- deriving functional requirements from privacy requirements,
- deriving non-functional requirements from privacy requirements, and
- identifying system components and processes requiring data minimisation.

Here, we select controls based on requirements. For example, if consent of the data subject is needed, mechanisms for obtaining and revoking consent, including transparency and intervenability capabilities, must be designed. If the impact assessment indicates higher risks to certain data, PETs may mitigate these risks while maintaining functionality. Otherwise, functionality must be reduced.

Relevant activities for architecture development include:

- selecting PETs for architecture candidates,
- choosing the most suitable architecture among candidates,
- compiling the list of privacy measures, and
- updating the impact assessments based on the architectural choices.

Stakeholder and operational considerations are transformed into technical specifications, and a system fulfilling the requirements is designed. Choices regarding PETs are also made during this stage. There might be multiple architectural candidates with or without PETs. To choose among them, establish evaluation criteria such as complexity, security objectives, privacy goals, legal compliance, development and maintenance costs, and conduct a comparative assessment. The architecture with the best score will be implemented.

Implementation, Testing, and Deployment. As systems are built, databases are integrated, interfaces are built, risks and impact must be assess. After testing, the system will be deployed and launched. Most importantly, during change management, privacy goals need to be adhered to. Changes must therefore be assessed to determine whether they elevate risks (and hence require additional privacy measures) or reduce them (making some measures no longer necessary).

Audits and Evaluations. Relevant activities include:

- verifying the achievement of privacy goals, and
- supporting the operation of existing PETs.

This is a critical step in the lifecycle as here legacy systems can be put on a path where they may be retrofitted with PETs. The first step is to recognise the desire for data protection by design. The second step involves reviewing the current status of the system (e.g., conducted impact assessments) and determining which privacy goals should be targeted in the next development cycle. The third objective is to allocate resources for the development of PETs.

Decommissioning. Relevant activities include:

- ensuring that decommissioning the system does not adversely impact organisational privacy goals, and
- updating the risk assessments of interfaced systems.

Sometimes a system has fulfilled its purpose, the organisation no longer has resources for its operation, or it has become obsolete and is replaced with a newer, more suitable system. In such cases, the old system is decommissioned.

4.3 Handling PETs in Procurements

A service provider procure work at any stage of the lifecycle. The procurer should specify requirements for minimising or avoiding data processing during this process. When procuring e-government services, consider the following:

- including privacy measures and PETs in the procurement specifications,
- requiring suppliers to demonstrate expertise in using and deploying PETs,
- establishing legal bases for supplier work in procurement contracts,
- asking the supplier to document the privacy measures of procured systems,
- collecting information on the supply chain of procured systems,
- assigning responsibility for residual risks (with contractual terms, if needed),
- establishing a protocol for reporting information security incidents.

5 Use Cases for PETs in Public Services

In this section we propose seven public sector use cases that benefit from the use of PETs. The use cases have been created based on the international case studies and results of interviews with 18 Estonian public sector organisations.

5.1 Secure Message Rooms for E-Government Services

On the Estonian X-Road development roadmap [15], message rooms are intended for providing services based on the data of an individual or an organisation. A message room is necessary when the data required to provide a service are not held by a single party and data need to be aggregated from multiple sources.

Using PETs, these services can be delivered so that sensitive data will not be gathered in an unprotected form in one central location. Data processing in the secure message room can be performed either entirely in an encrypted format or by utilising hardware security measures. The service can use the secure message room to link records of an individual and evaluate the specific business logic.

Figure 2 describes the model for a secure message used in e-government services. Examples that use secure message rooms include

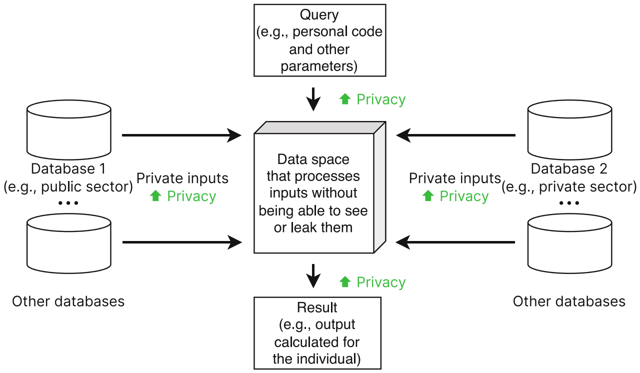

Fig. 2. Model for a secure message room for e-government services

- computation of welfare (e.g., energy subsidies) based on an individual's income and expenses,
- assigning personalised fines based on an individual's income,
- integrating private healthcare providers' data into the public health system,
- building privacy-preserving credit registries based on an individual's credit history at multiple lenders.

Advantages for Data Subjects. First, the service provider will not have to see all input records (unless allowed for quality or anomaly detection). Secondly, the data providers do not have to learn which of their records were used. Third, message rooms support transparency and intervenability, as the respective features can also be built using secure computing technology.

Advantages for the Government. The government can launch new services that may have previously been perceived as high-risk. These services will utilise public and private sector data and resources more efficiently. The privacy-preserving impact is significant for proactive, personalised, and event-driven services that require a high level of data integration.

Suitable PETs. Trusted execution environments, secure multi-party computation, and homomorphic encryption can be used for secure message rooms.

5.2 Privacy-Preserving Linking and Analysis Services

While secure message rooms are designed for building services, the linking and analysis service is meant for analytical use. It will be especially helpful in e-governments that deploy a distributed data architecture where each agency holds its own databases of people or company data. In such a setting, data-driven policy-making, research and machine learning model training may be complicated. The output in this scenario is not a computation result for an individual but rather a statistical aggregation, a machine learning model, or a report.

Figure 3 describes the model for a privacy-preserving linking and analysis service. This service will support an e-government in the following:

1. data-driven research in support of policy-making (e.g., early detection of public health concerns based on health records and fitness data, predictive maintenance for public infrastructure),
2. detecting fraud and anomalies across multiple databases (e.g., welfare fraud, tax fraud, money laundering),
3. analytics on transactions provided using a secure message room,
4. linking and analysis of public sector data before de-identified release to researchers or as open data.

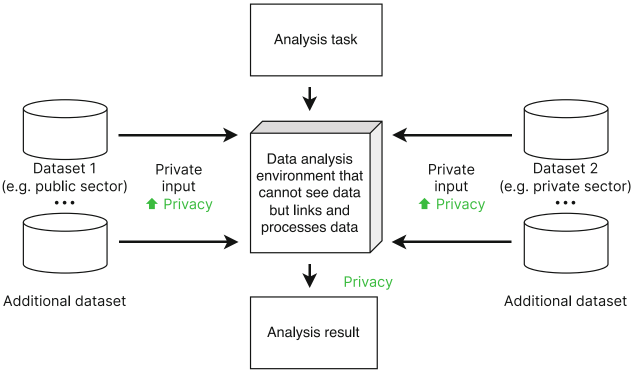

Fig. 3. Model for a privacy-preserving linking and analysis service

Advantages for Data Subjects. First, PET-based linking and analytics reduce processing of identifiable data in the public sector. Second, if a public sector authority links records from multiple databases, de-identifies them and releases to researchers or as open data, the process is more privacy-preserving than when researchers would perform the linking.

Advantages for the Government. First, data-driven policy-making ensures a better foundation for future policies. Second, including more public sector databases in fraud and anomaly detection will improve the quality of the services and also reduce loss of resources. Third, machine learning models for new innovative services (e.g., secure message rooms) can be trained using this service.

Suitable PETs. Trusted execution environments, secure multi-party computation, federated learning, homomorphic encryption, output privacy techniques like differential privacy or statistical de-identification can be used for this service.

5.3 Open Data Services

Open data services provide data to other services. This allows the creation of service hierarchies that rely on government data. PETs are not needed for data that are not directly identifiable (e.g., weather, public transport). However, it may be possible to include more data by using PETs, assuming that re-identification is infeasible. Figure 4 describes the model for open data services.

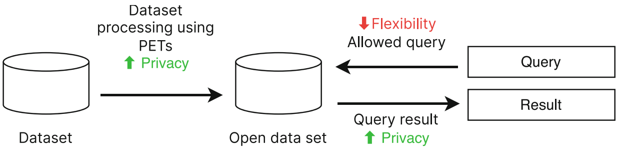

Fig. 4. Model for PET-based open data services

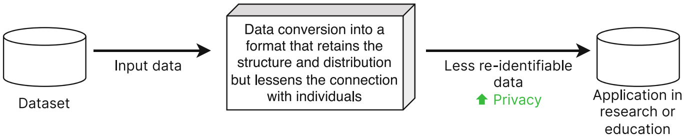
Dataset
Input data
Data conversion into a format that retains the structure and distribution but lessens the connection with individuals
Less re-identifiable data
⬆ Privacy
Application in research or education

Fig. 5. Model for a PET-supported data release service

Advantages for Data Subjects. Citizens will gain a greater awareness of state data. The use of PETs reduces the risk of re-identification of individuals in the datasets available through the services.

Advantages for the Government. More efficient consumption of public sector data would promote data-driven decision-making and improve management quality. For example, one could launch regionally focused services in education, healthcare, finance, transport, energy, and other sectors based on population statistics. These help municipalities in their planning.

Suitable PETs. The technologies need to be picked based on the sensitivity of the data. Query interfaces with restrictions will be sufficient for non-personal and non-confidential data. For more sensitive source data, techniques like differential privacy, synthetic data generation, aggregation and suppression should be used after sufficient risk analysis. Pseudonymisation is not suitable for protecting sensitive data in such services. Privacy-preserving linking and analysis can be used to process data before making them available in the API.

5.4 Publishing Databases as Open Data

For research and the development of new services there is a need for data that closely resemble real data but are de-identified. These activities (or PET use) are not often foreseen in the legal purposes of the respective databases. In such cases, it is practical to publish parts of a database after appropriate transformations.

The greatest risk associated with open data usage is the lack of control over their later re-use. The recipient of open data may re-identify individuals using any additional databases they have access to. Therefore, re-identification risk assessment for open data must be more comprehensive than for open data APIs. There may be a legal basis to release a database to a limited number of users (e.g., for research). This is not open data, but PETs can still be used to secure the processing. Figure 5 describes the model for publishing open data.

Advantages for Data Subjects. The use of PETs reduces the re-identification risk for individuals in published open data. People and companies benefit from the development of new services, research results, and policies.

Advantages for the Government. Making high-quality databases available for developers accelerates the creation of new services and enhances their quality (e.g., by reducing bias). Researchers have more source data to study. Using robust PETs reduces the re-identification based threat to an institution's reputation.

Suitable PETs. The careful application of synthetic data generation, differential privacy or aggregations may make re-identification infeasible for open data. Technologies for privacy-preserving linking and analysis can be used to process data before making them available. Pseudonymisation is not suitable for de-identifying sensitive data for publishing. For research and education, analyst workplaces with strict non-disclosure agreements may also be a solution.

5.5 Synthetic Digital Twins for Public Sector Data

E-government services need to be developed and maintained. Ensuring the quality of these services requires that the performance, correctness, and usability of the systems must be tested. Data protection regulations prohibit testing systems with live (or other kinds of personal) data. However, testing with purely random data is insufficient as it may not reveal edge cases and errors. In an e-government system that heavily relies on services, synthesised databases alone are not sufficient. There is also a need for services that use these data. Thus, a synthetic digital twin of the core services, embedded with synthesised data, is required to test larger systems. Figure 6 describes the model for a synthetic digital twin.

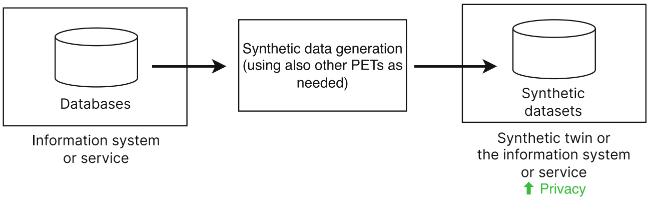
Databases
Information system or service
Synthetic data generation (using also other PETs as needed)
Synthetic datasets
Synthetic twin or the information system or service
⬆ Privacy

Fig. 6. Model for a synthetic digital twin for e-government data

Advantages for Data Subjects. The quality of e-government services increases.

Advantages for the Government. The development of new services in data-heavy domains like healthcare, finance, social welfare is accelerated.

Suitable PETs. Synthetic digital twins use synthetic data generation. However, such generation requires the analysis of source data and the creation of synthesis models. This implies the need for processing the source data. Such processing can be conducted using other PETs.

5.6 Privacy-Preserving Event Logging and Analysis

Logging plays a crucial role in digital services. Logs are used to detect errors, anomalies and misuse. However, a system log might also be a very sensitive database. Consider, for instance, logs of digital identity usage. When a user authenticates, digitally signs a document, or presents some evidence, a query is made to a digital identity validity verification service. Such service providers, when storing network connection information, can infer which services an individual uses and how often. Depending on the nature of the service (e.g., a health advisory system), such a log might contain highly sensitive personally identifiable information. The European Union's digital identity regulation is expected to impose restrictions on keeping such logs. Figure 7 describes the model for a private logging and log analysis service.

Advantages for Data Subjects. Personal data use in logging and log analysis is minimised, reducing the risk of personal profiling.

Advantages for the Government. Identity-driven e-government services have anomaly and fraud detection features that support maintenance and quality.

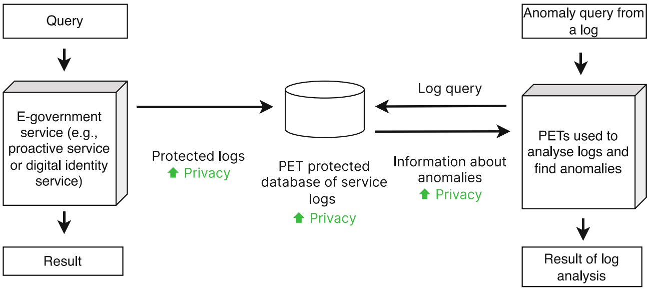
Query
E-government service (e.g., proactive service or digital identity service)
Result
Protected logs
⬆ Privacy
PET protected database of service logs
⬆ Privacy
Log query
Information about anomalies
⬆ Privacy
Anomaly query from a log
PETs used to analyse logs and find anomalies
Result of log analysis

Fig. 7. Model for a private logging and log analysis service

Suitable PETs. Trusted execution environments, secure multi-party computation, or homomorphic encryption can be used to build privacy-preserving event logging and analysis systems.

5.7 Proving Identities, Attributes or Data

Future digital wallets and digital identity may allow an individual to prove that they qualify for a service (e.g., based on age or a group membership). Today, such proofs are conducted by presenting the relevant source data. With an increase in such transactions, data can be amassed by the recipients of such proofs. An attribute proving service will minimise this processing of personal data.

The ability is not restricted to proving single values or their ranges. We can use zero-knowledge proofs to convince a verifier of a statement without providing all the source data, but just a cryptographic proof. For example, we can create a proof that an individual's health record meets a certain job-related health standard, an organisation's bank transactions comply with their funding rules, or a vehicle has travelled a certain distance in a certain regions (for subsidy or taxation purposes). Figure 8 describes the model for an attribute proving service.

Advantages for Data Subjects. The processing of personal data in digital transactions is minimised as personal data will not be transmitted to the verifier. As the prover (individual) compiles the proof using their own devices, they can directly view data about themselves, resulting in enhanced transparency.

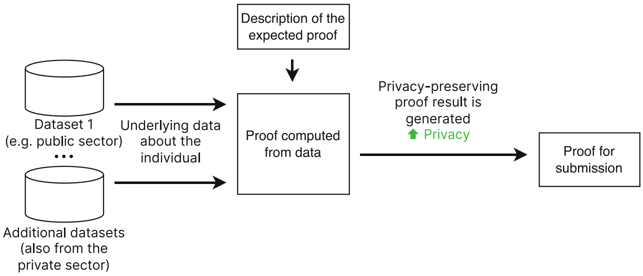
Description of the expected proof
Dataset 1 (e.g. public sector)
•••
Additional datasets (also from the private sector)
Underlying data about the individual
Proof computed from data
Privacy-preserving proof result is generated
⬆ Privacy
Proof for submission

Fig. 8. Model for an attribute proving service

Advantages for the Service Provider. The verification is partially moved outside the government and automated, reducing costs of standardised verifications.

Suitable PETs. Depending on the nature of the transactions, zero-knowledge proofs, blind signatures, and group and ring signatures can be used.

6 Conclusion

The concept [4] and roadmap [5] were published by the Estonian Ministry of Economic Affairs as a study of its artificial intelligence hub in 2023. No further activities as a part of the program have been announced at the time of preparing this paper. Further work by the authors of the strategy includes performing a deeper survey of international initiatives, monitoring implementation activities and keeping the PET catalogue up to date.

References

1. Article 29 Data Protection Working Party: Opinion 05/2014 on anonymisation techniques (2014). https://www.cnil.fr/sites/default/files/atoms/files/88197.pdf
2. Bogdanov, D., Kamm, L., Kubo, B., Rebane, R., Sokk, V., Talviste, R.: Students and taxes: a privacy-preserving study using secure computation. PoPETs **2016**(3), 117–135 (2016). https://doi.org/10.1515/popets-2016-0019
3. CNIL: Privacy Impact Assessment (PIA) (2018). https://www.cnil.fr/en/PIA-privacy-impact-assessment-en
4. Cybernetica AS: Privacy Enhancing Technology Concept (in Estonian). Technical report, Ministry of Economic Affairs and Communications (2023). https://www.kratid.ee/_files/ugd/980182_f1288bebbb57466ead0241748d49d8ec.pdf
5. Cybernetica AS: Privacy Enhancing Technology Roadmap (in Estonian). Technical report, Ministry of Economic Affairs and Communications (2023). https://www.kratid.ee/_files/ugd/980182_64478f7163b74f299f5879b6eea856af.pdf
6. Department of Economic and Social Affairs: E-Government Survey 2022. The Future of Digital Government. Technical report, United Nations (2022)
7. Fair Lines America Foundation: Declaration of John M. Abowd. https://s3.documentcloud.org/documents/21018464/fair-lines-america-foundation-july-26-2021-declaration-of-john-m-abowd.pdf
8. H.R.1565 - Student Right to Know Before You Go Act of 2019 (2019). https://www.congress.gov/bill/116th-congress/house-bill/1565/text
9. Infocomm Media Development Authority: Privacy Enhancing Technologies (PET) Sandbox (2022). https://www.imda.gov.sg/How-We-Can-Help/Data-Innovation/Privacy-Enhancing-Technologies-Sandbox
10. ISO/IEC TR 27550:2019 Information technology—Security techniques—Privacy engineering for system life cycle processes (2019). https://www.iso.org/standard/72024.html
11. IVXV raamistiku nõuded krüptosüsteemile. https://www.valimised.ee/et/e-haaletamine/dokumendid
12. McKenna, L.: Disclosure Avoidance Techniques Used for the 1970 through 2010 Decennial Censuses of Population and Housing. https://ideas.repec.org/p/cen/wpaper/18-47.html
13. Meijaard, Y., van Heesch, M., Cramer, R., Groenland, J.: Netherlands Cryptoland. Starting point of the cryptocommunications roadmap (2021). https://dcypher.nl/file/download/f67b5ad2-beee-4fdc-936c-5ac7b5fa3fea/netherlands-cryptoland-exploratory.pdf
14. Ministry of Economic Affairs and Communications: Estonian Digital Agenda 2030. Tech. rep., Ministry of Economic Affairs and Communications (2021). https://www.mkm.ee/en/e-state-and-connectivity/digital-agenda-2030
15. Nordic Institute For Interoperability Solutions: X-Road® Development Roadmap. https://x-road.global/development-roadmap
16. Raisaro, J.L., et al.: MedCo: enabling secure and privacy-preserving exploration of distributed clinical and genomic data. IEEE/ACM Trans. Comput. Biol. Bioinf. **16**(4), 1328–1341 (2018)
17. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/eli/reg/2016/679/oj

18. Rocher, L., Hendrickx, J.M., De Montjoye, Y.A.: Estimating the success of re-identifications in incomplete datasets using generative models. Nat. Commun. **10**(1), 1–9 (2019)
19. van der Sangen, M.: CBS explores possible privacy preserving techniques (2021). https://www.cbs.nl/en-gb/corporate/2021/14/cbs-explores-possible-privacy-preserving-techniques
20. Stevens, H., Haines, M.B.: TraceTogether: pandemic response, democracy, and technology. East Asian Sci. Technol. Soc.: Int. J. **14**(3), 523–532 (2020). https://doi.org/10.1215/18752160-8698301
21. Swiss Post: Cryptographic Primitives of the Swiss Post Voting System. https://gitlab.com/swisspost-evoting/crypto-primitives/crypto-primitives/-/blob/master/Crypto-Primitives-Specification.pdf
22. Technology Analysis Division of the Office of the Privacy Commissioner of Canada: Privacy enhancing technologies – a review of tools and techniques (2017). https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2017/pet_201711
23. The Royal Society: from privacy to partnership: the role of privacy enhancing technologies in data governance and collaborative analysis (2023). https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/From-Privacy-to-Partnership.pdf
24. The White House: US and UK launch innovation prize challenges in privacy-enhancing technologies to tackle financial crime and public health emergencies (2020). https://www.whitehouse.gov/ostp/news-updates/2022/07/20/u-s-and-u-k-launch-innovation-prize-challenges-in-privacy-enhancing-technologies-to-tackle-financial-crime-and-public-health-emergencies
25. Troncoso, C., et al.: Deploying decentralized, privacy-preserving proximity tracing. Commun. ACM **65**(9), 48–57 (2022). https://doi.org/10.1145/3524107
26. United Nations Committee of Experts on Big Data and Data Science for Official Statistics: United Nations Guide on Privacy-Enhancing Technologies for Official Statistics (2023). https://unstats.un.org/bigdata/task-teams/privacy/guide/2023_UN%20PET%20Guide.pdf
27. Weitenberg, E.: Secure and private statistics with distributed Paillier (2021). https://medium.com/applied-mpc/secure-and-private-statistics-with-distributed-paillier-8a186410b5af