



Bern, 12 February 2025

Report on the application of privacy preserving data science

Technical confidentiality management solutions for the joint use of data within the Federal Administration

Dossier number: 031.21-688/6/7/1



Contents

1	Introduction	4
2	Key concepts	6
2.1	Fundamentals.....	6
2.2	Input privacy and output privacy	9
3	Privacy enhancing technologies	13
3.1	Trusted execution environment.....	13
3.2	Homomorphic encryption	14
3.3	Secure multiparty computation.....	15
3.4	Distributed learning	16
3.5	Zero-knowledge proof	17
3.6	Differential privacy	18
3.7	Synthetic data.....	19
4	Use case 1: Platform for the confidential use of data science without data transfer	21
4.1	Context and objectives	21
4.2	Principle.....	23
4.3	Limits and challenges.....	25
5	Use case 2: Secure data enrichment based on private linkage	27
5.1	Context and objectives	27
5.2	Principle.....	29
5.3	Limits and challenges.....	32
6	Legal and governance framework.....	33
6.1	Processing of identifying data – review of the legal framework	33
6.2	Limits of the existing legal framework	35
6.2.1	Current status of attacks aimed at extracting identifying data: anonymisation no longer effective	35
6.2.2	Consequences for the legislation governing the use of data held by government authorities	36
6.3	The Five Safes: A governance framework for privacy preserving data science	36
7	Recommendation for the Federal Administration	40

Executive summary

The Federal Council has instructed the FDHA to draw up a concept for the application of privacy preserving data science within the Federal Administration based on the use of privacy enhancing technologies (PETs). These technologies open up new opportunities to create knowledge from the data held by the administration while simultaneously complying with data protection requirements. They allow data science techniques to be applied to the data of individuals without these being viewed or transferred, while also controlling the risk of disclosure associated with sharing the results. To keep up with the ever-growing demand for the delivery and linkage of the data held by the Federal Administration, the Federal Statistical Office (FSO) – through the Data Science Competence Center (DSCC) – is evaluating the potential of these technologies for simplifying, accelerating and, above all, limiting the administrative and technical burden involved in supplying these services while also guaranteeing the highest standards of privacy.

This report is aligned with several international initiatives, including a US presidential decree on the safe and secure use of artificial intelligence (AI) and the UN's work on privacy enhancing technologies in partnership with international providers of statistics. To illustrate the impact of these technologies on the joint use of data held by the Federal Administration, this report focuses on two standard use cases with applications that are relevant for all of Swiss society. The first describes the principle, and offers an implementation method for a platform to permit the confidential and secure application of data science without transferring data by providing remote access. This use case implements the principle of privacy by design by offering a standardised, automated technical solution, replacing the need for data delivery. However, the full potential of the data held by the administration can only be realised with the addition of the second use case, which makes it possible to pool several sources of data for enrichment purposes. This confidential data linkage service offers a modern, resource-efficient alternative to the type of linkage that is used today. The use of privacy enhancing technologies ensures that the data of individuals is protected but still available for joint use. These use cases pave the way for the more efficient and larger-scale secondary use of data both between federal offices and with external stakeholders such as academic researchers, (inter)governmental and non-governmental organisations and even the private sector, while also preserving and protecting privacy.

However, an examination of the current legal framework in the face of technical advances, particularly the boom in privacy enhancing technologies, reveals a number of limitations, not least the technical impossibility of guaranteeing that anonymisation is definitely irreversible. Furthermore, based on a strict implementation of the principle of purpose limitation, the secondary use of data is generally permitted to federal bodies only for purposes not related to specific persons, such as statistics, planning and research. Against this backdrop, privacy enhancing technologies offer a new set of tools for ensuring that privacy is preserved in a broader context than is covered by the current legal framework. However, this expansion calls for the creation of a risk-based governance framework to ensure respect for the rule of law as a whole. More extensive secondary use of data is then likely to be politically and socially acceptable, as this will ensure that the advantages it brings do not come at the expense of Switzerland's citizens, households and companies.

Privacy enhancing technologies therefore offer the Federal Administration new opportunities to improve efficiency and create value for the public good. As a result, the Federal Administration must ensure that it has at its disposal the IT infrastructure and technical expertise required to use these technologies. Combined with a legal framework and adequate governance, these technologies will play a key role in enabling the Federal Administration to fulfil its mandate ever more effectively for the benefit of Swiss society as a whole, while at the same time adhering to the constitution. Their implementation must support the Federal Administration in the fulfilment of its mandate without affecting existing ones, in particular those of the military administration and the Federal Intelligence Service (FIS).

1 Introduction

As part of the implementation of the federal data science strategy, the Federal Council has instructed the FDHA (FSO) to draw up a concept describing technical confidentiality management solutions for the joint use of data within the Federal Administration and prepare this report on the application of privacy preserving data science (point 6 of the Federal Council decision of 2 December 2022; EXE 2022.2746).

In order to respond to the challenges currently facing Swiss society (in areas such as health, climate, mobility and the environment) and to make informed decisions, government authorities may increasingly seek to develop public policy on the basis of this data. However, they can only achieve this systematically by fully exploiting the potential of the data collected by different administrative units. This data stands out thanks to its quality, its completeness and above all its reliability, and thus offers enormous potential for generating added value for the public good. At present, however its use is limited. Cross-cutting collaboration needs to be simplified both within and outside the Federal Administration without compromising the legitimate data protection concerns of Swiss individuals, companies and households. The Federal Administration must therefore develop privacy preserving data science practices that ensure respect for the rule of law when using the data entrusted to it.

The implementation of framework conditions enabling full use of the data held by the administration, including methods from the field of AI, would benefit both Switzerland's economy and its society. It echoes the Federal Council's desire to promote trustworthy data spaces based on digital self-determination in recognition of the fact that "the interconnection of data offers great potential for society, the economy, science and administrations".¹ Motion 22.3890 SECC-S of 22 August 2022 calling for the development of framework legislation for the secondary use of data² also underlines the need to find solutions that allow all data to be used (and not just open data).

At an international level, the presidential decree "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"³ recently issued by the Biden administration in the United States calls for the development of framework conditions to mitigate the risk of US citizens' personal data being disclosed by AI systems. In Europe, the Input Privacy-Preservation Techniques Project⁴ conducted by the United Nations Economic Commission for Europe (UNECE) illustrates the importance of privacy enhancing technologies for statistical organisations. In particular, the United Kingdom's Office for National Statistics is evaluating the possibility of using these technologies to protect mortality data.⁵

As part of the federal data science strategy, the report "Science des données pour le bien commun - cas d'application dans l'élaboration des politiques" [Data science for the public good – Use cases in policy development]⁶ showed that fully exploiting the potential of the data is a complex challenge that requires the involvement of multiple stakeholders and the development of numerous types of expertise

¹ <https://www.admin.ch/gov/en/accueil/documentation/communiques.msg-id-87780.html>

² <https://www.parlament.ch/en/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20223890>

³ <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence>

⁴ <https://unece.org/sites/default/files/2021-11/November%20workshop%202021%20HLG-MOS-PPT%20%281%29.pdf>

⁵ <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/applyingdifferentialprivacyprotectiontoonsmortalitydatapilotstudy>

⁶ [Science des données pour le bien commun - cas d'application dans l'élaboration des politiques - | Publication](#)

(organisational, ethical, legal, technical). The core issue associated with the application of data science is the need to establish a data governance framework that covers the entire data lifecycle (creation, collection, storage, security, access, deletion), focusing on three main areas:

- 1) IT infrastructure: This is addressed specifically in connection with the development of a concept for a collaborative data science platform (point 4 of the Federal Council decision of 2 December 2022; EXE 2022.2746) and forms the subject of a separate report.
- 2) The management of datasets: Through the national data management programme (NaDB; EXE 2023.1946),⁷ the Federal Council is promoting the multiple use of data over the long term through the 'once-only' principle, supported by initiatives for harmonising and standardising metadata such as the I14Y interoperability platform⁸ implemented by the FDHA (FSO) and Linked Data Services LINDAS operated by the FDHA (SFA).⁹ These initiatives, which catalogue the data collected by the administration and ensure its interoperability, are key to unlocking the full potential of the data held by the public administration. In addition to clearly defining the data held by the Federal Administration, metadata is also essential for the application of privacy preserving data science.
- 3) Confidentiality management: To satisfy the public's expectations with regard to the preservation of privacy, use of data held by the Federal Administration is at present largely restricted either by a strict ban on the secondary use of data for purposes other than that for which it was collected, or by an anonymisation obligation, particularly when data is published as open government data. In view of the scientific advances made over the last twenty years, particularly in relation to attacks that allow hackers to extract personal information from statistical publications or artificial intelligence models, anonymisation unfortunately appears to be a technically unachievable promise. In order to retain public trust in government authorities and permit the secondary use of the data they hold, the data privacy framework needs to be overhauled. New legal instruments backed up by privacy enhancing technologies need to be implemented to ensure that all secondary use of data held by the Federal Administration takes place within the 'Five Safes' framework¹⁰ (safe projects, safe data, safe people, safe settings, safe outputs).

Many stakeholders in both the IT sector and the legislative field must therefore work together in order to realise the full potential of the data held by the Federal Administration for the benefit of Swiss society as a whole. This report focuses on describing technical solutions for the implementation of a governance framework for confidentiality management. It also contains examples of practical applications for operationalising confidentiality management in the context of data governance. In particular, it explains the potential of privacy enhancing technologies (PETs) to create new opportunities for using data while also ensuring its confidentiality. Among other things, privacy enhancing technologies encourage the collaborative analysis of data that would otherwise be too sensitive to be shared with third parties or reused for purposes other than that for which it was originally collected.

In 2018, the UN Committee of Experts on Big Data and Data Science for Official Statistics (UN-CEBD) created a working group to support national statistical offices in using privacy enhancing technologies

⁷ <https://www.bfs.admin.ch/bfs/en/home/nadb/nadb.html>

⁸ <https://www.i14y.admin.ch/en/home>

⁹ <https://lindas.admin.ch/>

¹⁰ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>

(UN Task Team on Privacy-Enhancing Technologies), which has published a guide on the topic.¹¹ Confidentiality management, and more specifically controlling the risk of disclosure, is a key topic for statistical offices, which must guarantee the confidentiality of the data entrusted to them in order to comply with Principle 5 of the Code of Good Practice for Official Statistics.¹² The OECD, meanwhile, has published a report that is intended to give policy makers a better understanding of how these privacy enhancing technologies can be used to enhance privacy while also allowing data and information to be shared.¹³ The Biden administration's executive order mentioned above also clearly identifies privacy enhancing technologies as a key technical tool for the safe and effective use of data within the Federal Administration.

Against this backdrop, this report aims to put forward technical concepts capable of having the greatest possible impact at federal, national and international level. These concepts therefore need to be portable, exportable and reusable, and compatible with initiatives to standardise metadata through the I14Y platform and promote a common vocabulary thanks to the ontologies proposed by Linked Data Services LINDAS.

After defining the fundamentals and the relevant stakeholders associated with the joint use of identifying data, the report presents various privacy enhancing technologies and sets out their respective advantages and weaknesses for use within the Federal Administration.

It goes on to present two use cases before reviewing the existing legal basis and explaining how privacy enhancing technologies can help implement the fundamental principles of privacy. Finally, it proposes recommendations for the implementation of a governance framework for confidentiality management that makes the best possible use of the technologies presented in this report while also adhering to the legal framework for data protection. This governance framework must ensure that the technologies support the Federal Administration in the fulfilment of its mandate, but without affecting existing ones, in particular those of the military administration and the FIS.

2 Key concepts

2.1 Fundamentals

Before presenting the different technologies that permit the implementation of a data science methodology for the joint use of data within the Federal Administration in detail, we need to define the concepts used in this report. Data protection is by its very nature an interdisciplinary undertaking that is also largely dependent on the context in which data is used. As such, the meaning of the terms used in this report need to be clarified. The definitions chosen are determined firstly by the use case we are examining, namely the joint use of data held by the Federal Administration, and secondly by the need to follow the standard language of English and make it easy to link the concepts to the existing scientific literature.

In this report, an **identifiable unit**, also referred to as a **privacy unit**, is any natural person or legal entity, or any grouping thereof such as a household, that generates information that is not in the public domain and whose content must be protected. In the case of natural persons, an identifiable unit is also referred to as a data subject. The generalised approach adopted in this report is necessary because, from a

¹¹ <https://unstats.un.org/bigdata/task-teams/privacy/guide/>

¹² <https://unstats.un.org/unsd/dnss/docs-ngaf/National%20Code%20of%20Good%20Practice.pdf>

¹³ https://www.oecd-ilibrary.org/science-and-technology/emerging-privacy-enhancing-technologies_bf121be4-en

technical point of view, ensuring the protection of a group of individuals such as a household requires more sophisticated measures than ensuring the protection of each natural person within the group.

The term **identifying data** therefore refers to all the information collected about either one identifiable person or a group thereof. Since the nature of a piece of data does not determine its identifying characteristics, all data associated with an identifiable unit must be considered identifying. Modern re-identification techniques are based not on the value of individually captured pieces of data but on their combination. This is the principle behind linkage attacks, which are described in more detail in section 6.2.

If the identifiable unit is a natural person, in line with data protection legislation such as the Federal Act on Data Protection ([Art. 5 let. a FADP](#); [SR 235.1](#)) we then refer to **personal data**. In the case of legal entities, the legislation uses the term **legal entities' data** (Government and Administration Organisation Act (GAOA); [SR 172.010](#)). The application of privacy preserving data science extends beyond the framework of these laws and also covers, for example, the Ordinance on the Conduct of Federal Statistical Surveys ([SR 431.012.1](#)), which states that “the results of the surveys shall be made accessible to the public in a form that precludes any identification of the persons, households, companies or establishments surveyed”. This example, which includes the protection of natural persons, legal entities and groupings thereof, informs the decision to make a distinction in this report between personal data and identifying data.

A **data product** is any product that is derived from identifying data and developed through the application of data science. These products generally fall into three main categories:

- **Microdata:** These are datasets providing information at the level of one or more data subjects (or privacy units).
- **Data analysis results:** This category includes the production of statistics, and more generally any information relating to a group of privacy units. These objective indicators are particularly useful for developing public policy.
- **Machine learning and artificial intelligence models:** These may in fact be trained on identifying data, with the risk that this data could be disclosed either during the training process itself or when the model is subsequently being used.

The application of privacy preserving data science therefore ensures that the identifying data used to generate these products is protected. In this context, a distinction needs to be made between two different objectives:

- **Privacy**, the aim of which is to ensure that data products comply with a set of standards. These are directly influenced not only by legislation but also by contemporary social and cultural reality, which determines what is and what isn't acceptable when a data product is created. For example, the concept of data sensitivity as it relates to privacy is contextual; it cannot and must not be anticipated. By way of illustration, a natural person's gender is often considered to be non-sensitive information, since it is common knowledge. However, making public the fact that the gender composition of a household has changed from one year to the next does constitute the disclosure of information that is considered sensitive, namely a change of gender. None of these specific cases can be anticipated when a data product is created, not least because the sensitivity of a piece of information is the result of a value judgement that depends on the state of a society at a given moment in its history.
- **Confidentiality**, which encompasses all measures aimed at ensuring that identifying data remains secret. The concept of confidentiality is therefore objective. It is future-proof and is much stronger and more difficult to implement from a technical perspective.

In this report, the objective of privacy preserving data science is to ensure the confidentiality of each privacy unit's identifying data, and not just to simply verify that products comply with current standards. The aim is to develop a formal framework that is independent of the evolution of standards and whose goals can be evaluated objectively. This is based on the position of the U.S. Census Bureau, which is legally bound to ensure the confidentiality of natural persons in all the data products it makes available.¹⁴

The confidentiality of identifying data must be ensured both while a data product is being created and when it is subsequently made available and used. As such, the first step is to ensure the **security of the data**, in other words to prevent unauthorised access to identifying data. The main tools used to ensure data security are based on cryptographic techniques, authentication and access controls. In a second step, confidentiality must ensure that no identifying data is **disclosed** when the product is made available, in other words that it is not made public when results are published voluntarily. These two aspects of confidentiality complement each other and define two categories of technology that will be described in the next section.

Finally, in order to place the objectives of privacy preserving data science within the context of the FADP (RS 235.1)¹⁵ and the General Data Protection Regulation (GDPR),¹⁶ it is important to specify the meaning of the following terms:

- **Anonymisation:** Process via which identifiable data is irreversibly modified in such a manner that it can no longer, or only via means that are not reasonably likely to be used, be directly or indirectly attributed to an identifiable person.
- **Pseudonymisation:** The modification of identifying data in such a manner that it can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the data is not attributed to an identifiable person.

It therefore seems clear that ensuring the confidentiality of identifying data is equivalent to the concept of anonymisation, and that any technique falling within the scope of pseudonymisation will thus be considered inadequate for guaranteeing confidentiality without additional organisational and/or contractual measures.

From a purely technical point of view, recent scientific advances show that neither anonymisation nor the absence of the risk of identifying data being disclosed can be guaranteed. The publication of a data product, of whatever type, always involves a non-zero risk of identifying data being disclosed. However, it is possible to control this risk with the aid of technical solutions. These technologies make it possible to navigate what is commonly referred to as the **privacy-utility trade-off**. This concept intuitively conveys the reality that it is not possible to publish a data product that exactly reflects the set of identifying data while also guaranteeing a high level of confidentiality. As such, the cost of significantly limiting the risk of disclosure is that the product becomes less useful, and vice versa. The privacy enhancing technologies presented in the rest of this section help quantify and therefore navigate this trade-off in an informed way. However, the choice of an acceptable level of disclosure risk and/or a minimum level of usefulness is a highly political and societal question that cannot be resolved through technological means.

¹⁴ <https://www.govinfo.gov/content/pkg/USCODE-2007-title13/pdf/USCODE-2007-title13.pdf>

¹⁵ <https://www.fedlex.admin.ch/eli/cc/2022/491/en>

¹⁶ <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

2.2 Input privacy and output privacy

The creation and distribution of a data product take place via a use chain for the identifying data that requires the involvement of the following stakeholders:

- The data subject(s): privacy unit that generates the identifying data.
- The input party: organisation that collects the identifying data and mandates its use through the creation of a data product. For example, a national statistical office producing statistics such as the poverty index, or a company collecting data from connected objects.
- The computing party: organisation that supplies the computing power required to create the data product. The computation is generally carried out on remote servers, in other words via a cloud infrastructure. This infrastructure may be operated either by a trusted organisation such as a federal office or by a private company that grants access to servers that are shared by all the service provider's clients and are thus classed as public.
- The result party: group of individuals to whom the data product is distributed. In the case of an administration, for example, any data product of the type 'open government data' is distributed to the public without restriction. This constitutes the largest distribution group for which confidentiality controls must be at their strictest, as the risk of identifying data being extracted is at its highest. The dissemination of a product may also be limited to a more restricted group, such as employees of the Federal Administration, who are all bound by official secrecy and for whom the risk of an attempt to extract identifying data is lower than in cases where a product is disseminated to the general public.

Privacy enhancing technologies (PETs) are a set of technical solutions for controlling the risk of a loss of confidentiality with respect to identifying data along its entire value chain according to the trust placed in the various stakeholders. The term 'trust' should be understood here in the broadest sense of the word, in other words any reason that explains the non-disclosure of identifying data to another party. This includes, for example, the existence of a law prohibiting the exchange of identifying data between the parties, a lack of trust by data subjects towards the input party, or the absence of a legal framework or organisational measures to prevent any attempt to extract identifying data. The absence of a trust relationship between two or more stakeholders can be overcome through the use of privacy enhancing technologies, which can be divided into two main categories:

- Technologies ensuring **input privacy**, which seek to enable two or more of the stakeholders described above to submit data for computation without any other stakeholder being able to access original identifying data other than its own. These technologies are usually based on cryptographic techniques and therefore form part of the solutions ensuring data security.
- Technologies ensuring **output privacy**, which aim to limit the disclosure of identifying data from data products. In this context, the product is disseminated voluntarily by the input party. This process is also known as (statistical) disclosure control, a field that is very familiar to national statistical offices, who have been developing this practice for several years.

Depending on the trust relationships between the different stakeholders in the value chain of the identifying data, the organisation guaranteeing confidentiality, generally the input party, can therefore use a combination of technologies to ensure input and output privacy. Figures 1, 2 and 3 provide examples of trust schemas, which are used to identify those points in the use chain for identifying data where technologies to ensure input/output privacy can be used to overcome a lack of trust between parties.

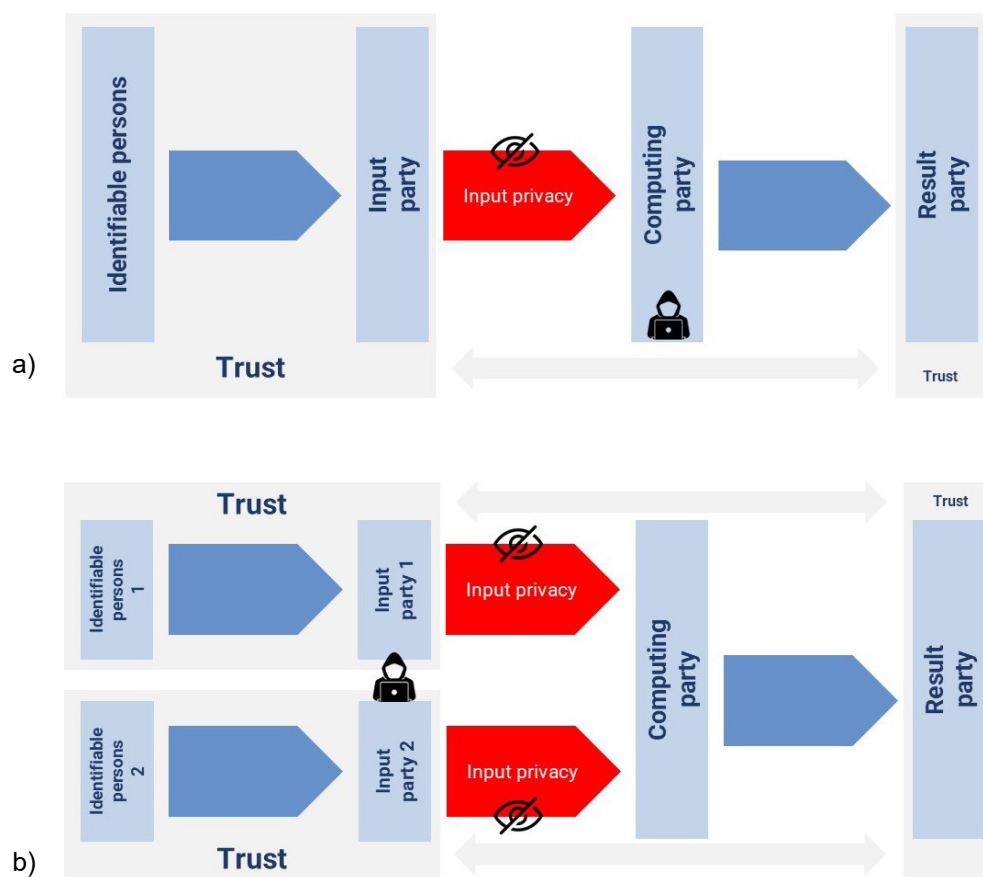


Figure 1: Sample trust schema for the use of technologies to enhance input privacy with a) a single input party and b) two input parties. In both cases the input parties and the result party are linked by a trust relationship

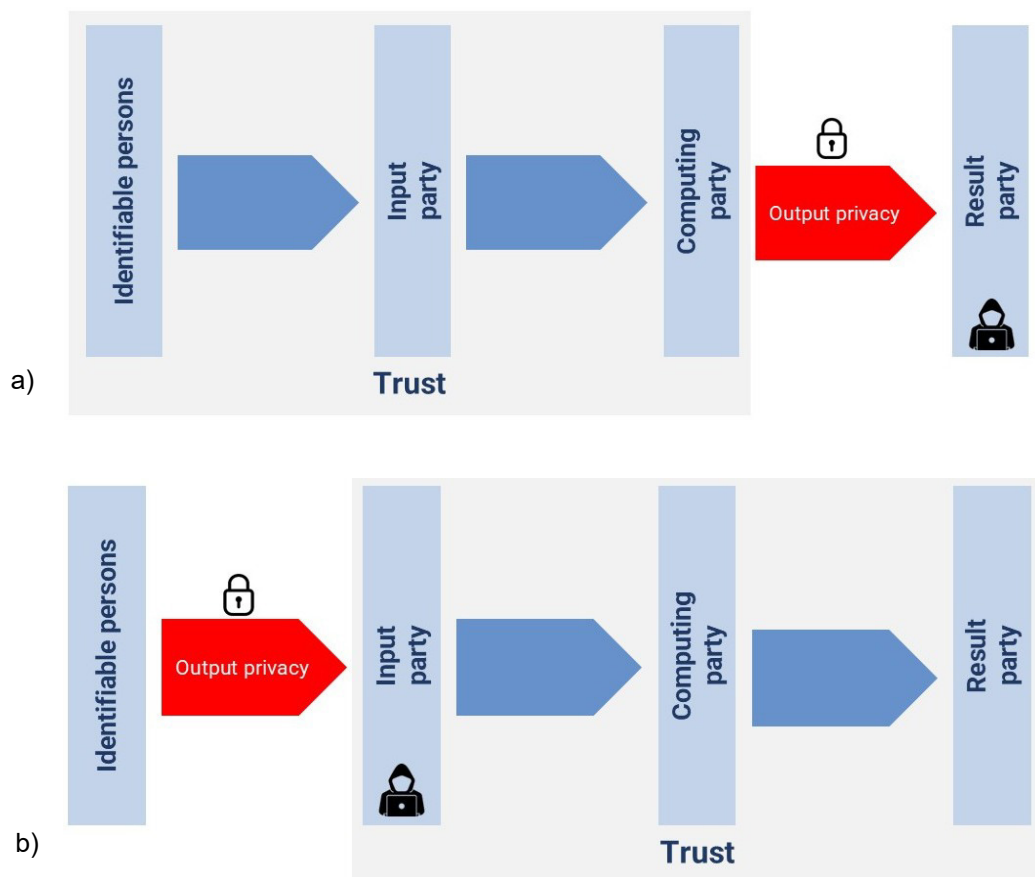


Figure 2: Sample trust schema for the use of technologies to enhance output privacy with a) the reference schema for open government data and b) the schema when the data subjects (identifiable units) do not trust the input party

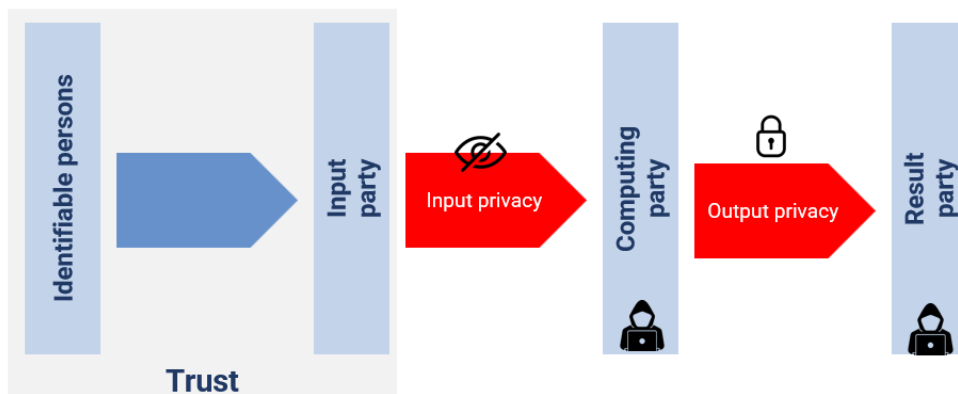


Figure 3: Sample trust schema necessitating a combination of technologies to enhance both input privacy and output privacy

3 Privacy enhancing technologies

3.1 Trusted execution environment

Explanation of principle

The aim of a trusted execution environment (TEE), also referred to as confidential computing, is to guarantee the security and integrity of code and data against software and hardware attacks during the execution of a computer program. This technology permits the private execution of computer programs on remote infrastructure without the need for the input party to be a trusted party. The TEE ensures the integrity of the code and protects the data against any manipulation within the environment by preventing all unauthorised external access. The integrity of the executed code is guaranteed by a certificate of execution generated by the input party through a cryptographic proof of authenticity. The implementation of these environments is based on hardware components that are supplied by certain IT manufacturers or service providers and are generally integrated into specific processor models.

Limitations and risks

A TEE is directly linked to the processor models on which it can be deployed, with certain implementations sometimes only available via a specific cloud service, thus creating a dependency with regard to the chosen provider. Certain TEE models are also limited in terms of the computational memory available. While some implementations make it possible to circumvent this limit, excessive usage of this alternative is often considered to be bad practice.

In addition, since the technology is directly linked to the hardware, it is exposed to technical vulnerabilities, whether due to direct access to the machine or to the subsequent discovery of security flaws rendering the system obsolete and impossible to update. Its functionality also varies significantly depending on the provider and the underlying technology, adding a further layer of complexity to the deployment of a TEE.

In a TEE, the code being executed is not necessarily secret, as it may be known to the system administrator. While the environment offers protection during execution of the code, the resilience of the code itself to external attacks (via the interfaces that it supplies) may remain inadequate. This situation potentially exposes the system to vulnerabilities with respect to side channel attacks.

Promises and maturity

TEE technology has already been adopted in various sectors, notably to safeguard the execution of sensitive code on smartphones (classed as untrusted devices). Financial services applications frequently make use of this technology to ensure the security of payments and identity checks. One of the major advantages of this technology is its ability to integrate with other PETs without significantly compromising the performance of code execution.

Numerous providers of cloud services (Azure, AWS, etc.) also offer this technology within their platforms, making it easier to integrate within organisations. The principal focus of current research is on identifying the risks associated with the different potential attacks on this technology and on developing a standard, something that is still lacking in this area. However, certain players such as Globalplatform and Enarx are making efforts to remedy this.

One of the key benefits offered by certain TEEs compared with other PETs is the flexibility to port existing algorithms (such as those used for data analysis) to the secure environment without any drastic modification of the source code. This is highly dependent on the framework used when implementing a TEE and is an area of research that is still in full expansion.

A TEE can have various applications in data science. For example, a provider (owner of a learning model) may make its model available on a TEE so that a client (data owner) may use it while ensuring

that its data is sent there securely. This means that the client does not have direct access to the model, while the provider does not have direct access to the data.

3.2 Homomorphic encryption

Explanation of principle

Homomorphic encryption is a cryptography technique that allows computations to be performed directly on encrypted data thanks to its mathematical properties. This means it is not necessary to decrypt the data in order to carry out statistical analysis, hypothesis testing or AI inference.

For example, this technique allows a provider of sensitive data (A) to outsource the processing and storage of identifying data to (B), a computing party in a cloud infrastructure. Now let's consider the case where A is an employee of a federal office who wishes to use the machine learning model made available by B. This model predicts fraud risk based on identifying data in tax returns. A homomorphically encrypts the data and transmits it to B, who runs the encrypted data through the model and then sends the results – also encrypted – back to A. Finally, A decrypts the results to find out the inferred fraud risk, while B has learned nothing about the data.

At the time this report was written, homomorphic encryption was the only technology allowing data to remain encrypted throughout the entire process, thus eliminating the need to trust the computing party, which never has access to decrypted data.

Fully homomorphic encryption (FHE) is the most advanced form of homomorphic encryption. It allows an unlimited number of mathematical operations to be performed on the encrypted data and therefore opens up more complex applications such as the training of deep neural networks.

Limitations and risks

Homomorphic encryption poses similar risks to other encryption techniques. Specifically, it requires secure storage of the encryption keys and resilience to brute force attacks, cryptanalysis, side channels and data interception.

The chief limitation of homomorphic encryption is the considerable amount of IT resources needed. Encrypted data is significantly larger, meaning that more storage space is required. Computations on encrypted data are also considerably slower than on plaintext data, making homomorphic encryption unsuitable for real-time applications.

It should also be noted that homomorphic encryption is an input privacy technology. As such, although the data remains encrypted with computing parties, there is no guarantee whatsoever with regard to the risk of disclosure once the data product has been decrypted by the result party.

Finally, and above all, homomorphic encryption only offers a limited number of analyses. Only linear operations can be computed exactly, so approximations have to be used for any more complex operations. Guaranteed encryption therefore comes at the expense of accuracy.

Promises and maturity

Homomorphic encryption is used successfully in practice, and libraries of open source code are available. One of these use cases is private set intersection, which allows parties to find common elements between two sets of identifying data belonging to different input parties who do not wish to share their data with each other. The provision of machine learning models, as in the fraud risk example described above, is another notable application.

Homomorphic encryption is an active field of research that has made considerable progress over the last five years, in particular through the development of specialised hardware solutions to reduce computation time and make the technology viable. If progress continues at a meaningful pace, homomorphic encryption will be available for use in multiple applications, opening up enormous potential in a wide range of fields.

3.3 Secure multiparty computation

Explanation of principle

Secure multiparty computation (MPC or SMPC) is a form of secure computation that enables a group to jointly perform computations without disclosing the participants' private inputs (data). MPC is a sub-branch of cryptography. Unlike traditional cryptography, in which the attacker is outside the system, MPC aims to guarantee the security of participants' data vis-à-vis the other participants. The participants agree on the algorithm to apply, then use an MPC protocol to jointly compute the output of this algorithm based on their secret inputs, without revealing them. Each participant only learns what can be deduced from the output. MPC guarantees that each party's inputs remain private (input privacy) and ensures the integrity of the results obtained.

There are two main categories of MPC algorithm. The first allows all types of algorithms to be executed, in particular those that enable data products to be created. This category makes extensive use of the concept of secret sharing, whereby inputs are split into shares and distributed among the parties. The MPC algorithm then performs the computations on the separate shares, which are only recombined once execution has been completed in order to keep each party's inputs secret throughout the execution of the algorithm. The second category consists of algorithms that are dedicated solely to the creation of very specific data products (such as linkage for sets of private data) and are highly optimised for this purpose.

Limitations and risks

MPC protocols offer a higher level of security with a formal guarantee, but usually involve cryptographic operations that result in higher computation and communication costs. This technology is therefore only suitable for specific tasks.

There are a range of (more or less) strict threat models for MPC, including the possibility of corrupt parties or even collusion between them. For example, some algorithms are secure even if the majority of parties involved are corrupt. When using MPC, it is therefore important to define a clear threat model and exercise caution when selecting suitable algorithms and libraries.

The execution of existing algorithms often requires code to be rewritten to permit the transition to MPC. In addition, open source code software libraries for MPC frequently do not offer the same range of specific data science functionalities, resulting in extra work when adapting code for MPC.

Finally, while MPC limits each party's ability to access the other parties' input data, it does not guarantee the confidentiality of the computed result, which in some cases can reveal sensitive information. As a result, MPC requires careful consideration of the computations to be performed, and often needs to be combined with techniques such as differential privacy that are geared toward preserving output privacy.

Promises and maturity

MPC has applications in scenarios where input privacy is critical, such as collaborative analysis of data in healthcare, finance and other sensitive fields. For example, MPC techniques were used in Estonia to carry out a study on the correlation between working during university studies and academic failure. This study required data from two independent state entities who have no legal basis for sharing it. Without the help of MPC, therefore, this study would never have seen the light of day.

The vast majority of projects that use MPC techniques rely on proprietary software libraries. At present there are no open source code libraries for MPC that are sufficiently mature for production purposes. Current research is focused on optimising MPC protocols to mitigate the overload in terms of computation and communication. MPC should become more efficient as cryptographic techniques develop, making it more viable for a broader range of applications. Combining MPC and differential privacy (see below) as a means of preserving output privacy is also a promising area, but is currently still at the research stage.

3.4 Distributed learning

Explanation of principle

Distributed learning enables several parties to train a machine learning algorithm, more specifically a neural network, in a way that guarantees the security of the training data among the parties (input privacy). The data remains stored on its owner's infrastructure and is never exchanged between the parties. Only the model parameters, in other words the result of the learning, are shared between the parties. These techniques require an external computing party to coordinate the process: they never need access to the input data but play a key role in aggregating the training parameters and distributing them to the different parties.

There are two main distributed learning approaches: federated learning and split learning. In federated learning, the participants each have their own copy of the model, which they train individually on their private data. The participants then send the parameters of their model to a central server for aggregation. These are used to construct a global model that is redistributed to each participant for a new training iteration. In split learning, the model is split up so that each of the participants can train a separate sub-section of the model on their infrastructure using their own data. Each sub-section of the model produces an intermediate output. During the training process, the central server also trains a model that combines these intermediate outputs. In contrast to federated learning, split learning is performed sequentially. However, split learning offers better protection for input data by concealing more information on the model parameters and architecture from the other parties, which also limits their ability to examine the model.

Limitations and risks

Compared with the traditional approach of centralised training of machine learning models, distributed learning generates additional costs in terms of the communication between the parties and the central server. The learning time is also extended for techniques that require synchronisation between the parties at each stage of the model optimisation algorithm. In general, AI models trained by means of distributed learning perform less well than equivalent models trained on centralised data.

Distributed learning aims to guarantee the privacy of the private input data. However, the model parameters transmitted during the learning phase and the resulting model can reveal information relating to participants' data, which is why it must be combined with an output privacy technology in order to control the risk of disclosure.

Finally, distributed learning is also vulnerable to the corruption of one or more parties or of the aggregation server. For example, a participant might enter false training data in order to modify the behaviour of the model in a targeted manner and thus serve its own interests.

Promises and maturity

Distributed learning promotes collaboration between several organisations by enabling them to jointly train a model without having to centralise their data, which resolves certain problems associated with input privacy. Distributed learning also takes advantage of the diversity of the data sources to promote the development of more robust and efficient models. For example, health data is not always available in sufficient quantities to permit the training of efficient AI models for medical diagnostics. Distributed

learning allows researchers to train a robust AI model thanks to data from several hospitals, while also improving input privacy.

Several open source libraries are available that permit distributed learning of AI algorithms. These integrate well with traditional machine learning libraries, making this technology accessible.

3.5 Zero-knowledge proof

Explanation of principle

Zero-knowledge proof (ZKP) is a cryptographic protocol that allows a 'prover' to prove to a 'verifier' that a statement is true without revealing any auxiliary information other than its validity. For example, a ZKP can be used to verify the following proposition: a machine learning model for predicting solvency has been run on the entire financial history of company X, but without disclosing either the specific details of the model or the company's financial data. ZKPs are probabilistic, in other words there is always a residual risk that the statement may be considered to be false even though it is actually true. However, this risk can always be reduced arbitrarily by engaging additional computing capacity.

The main areas of application for ZKPs are authentication, identification and cryptocurrency. In data science, ZKPs can facilitate data sharing and collaboration between stakeholders in scenarios where identifying information must be protected. For example, they can be used to verify the nature of identifying data without revealing them directly, or to verify a person's identity without disclosing personal information. The data flow is therefore protected against leaks, and the stakeholders do not share their data with a potential computing party.

Limitations and risks

ZKP verification is a resource-hungry task, requiring either multiple communications between verifier and prover or a very large amount of computing capacity. As a result, the practical applications of ZKP are currently limited to very simple problems if it is to be usable either on a large scale or in real time.

Creating new ZKP protocols is often a theoretical challenge, as it is difficult to translate a human problem into a mathematical formula. Implementing them therefore requires not only IT developers but also people with a good knowledge of mathematics and cryptography. Going forward, many ZKP protocols will no longer be secure with the advent of quantum computing capacity.

As well as being an input privacy technology, ZKPs often require a certain level of trust between input parties and computing parties, meaning they are only applicable for very precise trust schemas.

Promises and maturity

The maturity of this technology varies. Some ZKP applications are already well assimilated and applied, while others are still at the research stage. Since they are necessary for the operation of blockchains, where transactions can be verified without sharing identifying data, ZKPs are benefiting from numerous efforts aimed at overcoming their current limitations, with work underway to develop a taxonomy and standards.

One typical example of a ZKP application in data science is verifying the correct execution of an algorithm without disclosing the data used. In machine learning and AI, the owner of a model can therefore verify the result of a model without having to share it. ZKPs also play a key role in the implementation of electronic voting systems, thus guaranteeing the validity of elections and the anonymity of electors. Finally, in February 2023 the European Union approved a 'digital identity' using ZKPs and enabling all European residents and citizens to identify themselves and verify personal information for dematerialised services such as hiring a bicycle or paying their taxes.

3.6 Differential privacy

Explanation of principle

Differential privacy refers not to a single tool or algorithm but to a methodological framework defining the mathematical properties that an algorithm for statistical analysis or machine learning must satisfy in order to quantify in mathematical terms the disclosure risk associated with sharing the associated data product. It is therefore an output privacy enhancing technology.

More specifically, differential privacy works by randomising the output of an algorithm so that any attempt to extract identifying data from the data product is rendered unreliable. An algorithm that satisfies the conditions for differential privacy guarantees precise control of the risk of identifying data being disclosed, irrespective of the auxiliary information and the technical means available to the attacker. This risk is quantified by a parameter known as the privacy-loss budget, the choice of which depends on the sensitivity of the data and the intended use of the protected data product.

Differential privacy can also be applied directly to the original dataset (referred to as local differential privacy). This approach aims to render the data unreliable by randomly perturbing it so that it can no longer be used against the identifiable unit to which it belongs. This mechanism, known as randomised response, has been used for several years by national statistical offices, polling organisations and more recently the private sector to guarantee the truthfulness of a response, or more simply to encourage acceptance of data collection, in relation to sensitive topics such as extra-marital relations or the telemetry (localisation, installed applications, keystroke logging, etc.) of personal computers.

Limitations and risks

Differential privacy mechanisms require the introduction of random perturbation into the algorithm. It is therefore necessary to evaluate and clearly communicate the framework within which the perturbed data products, or microdata, can be used and with what level of confidence. This evaluation is carried out on a case by case basis and requires advanced technical expertise in privacy and mathematical statistics.

There are currently numerous initiatives aimed at developing standards or simple guidelines for calibrating the scope of privacy based on the context and the degree of sensitivity of the data product.¹⁷ However, deciding on the privacy-loss budget and hence the adjustment of the privacy-utility trade-off remains a largely political challenge that calls for close discussion between differential privacy specialists, regulators and political decision-makers. There is no universal solution, as the choice of this parameter is dependent on the data privacy framework, taking all the Five Safes¹⁸ into account. Managing the expectations of the recipients, in other words the 'safe people', is particularly important, as the level of reliability of the published product can only be assured for very specific uses.

It should also be noted that differential privacy requires specific technical skills over and above mathematical and statistical expertise. Implementation is non-trivial in terms of IT infrastructure, particularly due to the need to ensure rigorous monitoring of the privacy-loss budget and of the versions of the data used in the analyses.

Promises and maturity

¹⁷ <https://www.nist.gov/news-events/news/2023/12/nist-offers-draft-guidance-evaluating-privacy-protection-technique-ai-era>

¹⁸ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>

Differential privacy, formally introduced in 2006,¹⁹ is the subject of intensive academic research and experimentation. It forms an integral part of ongoing work on statistical disclosure control and is a response to advances in knowledge relating to attacks aimed at disclosing identifying data, such as reconstruction, linkage and affiliation. Today, differential privacy is deployed in the productive environment both by private sector operators (Apple, Google, Microsoft) and by government institutions (U.S. Census Bureau). Open source code libraries with a sufficient level of maturity and reliability are therefore available. Since differential privacy algorithms are only very marginally more complex than traditional algorithms, the additional computation cost is not a limitation.

Differential privacy is an especially promising output privacy technology for the following reasons:

- A mathematically guaranteed limit on the risk of identifying data being extracted from a data product whose validity is assured, irrespective of the auxiliary information and technical means available to a party attempting to extract this information. This guarantee ensures that any attempt at extraction cannot be carried out in a reliable manner, while fixing the maximum level of reliability that a potential attacker can achieve.
- The cumulative risk of disclosure can be quantified even with successive, independent publications of different types of data products.
- The mechanism for adding noise to data can be communicated with the greatest possible transparency without compromising the privacy guarantee, allowing for an open public debate on the privacy-utility trade-off.
- The guaranteed control of the disclosure risk for identifying data that is offered by differential privacy can be used to permit the secondary use of data held by the Federal Administration, either within the latter or between the Federal Administration and the rest of Swiss society (researchers, private sector, civil society), without sharing the data, while controlling the risk to which the individuals contributing their data are exposed and assuring them that there is no reliable way of extracting any identifying data. Section 2 presents this use case in detail along with its great potential to benefit society.

3.7 Synthetic data

Explanation of principle

The aim of synthetic data generation is to create an alternative dataset with granularity at the level of one or more data subjects, but without obviously disclosing the identifying data of the original dataset. There are two types of synthetic dataset:

- Fully simulated or dummy data, which simply reproduces the characteristics of the original data. More specifically, this type of synthetic dataset merely shares the same metadata (variable name, variable types, possible values, etc.) as the identifying data, but only contains randomly generated information. This information is primarily used to test data analysis pipelines.
- Data generated in order to reproduce certain statistical properties of the original dataset. Generation is therefore not purely random, even if it remains partly stochastic. In this case, since the information on the dataset is shared, there is a non-zero risk of identifying data being disclosed that needs to be controlled.

Generators of the first category of synthetic data are usually based on simple random number generators. The second category uses more complex algorithms derived from applied mathematics and

¹⁹ Dwork, C. (2006). Differential Privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds) Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science, vol 4052. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11787006_1

algorithms from machine learning and artificial intelligence (random forests, neural networks, generative adversarial networks, etc.).

Limitations and risks

Each synthetic dataset is generated in order to reproduce certain properties of the original dataset. The nature and scope of these properties must therefore be clearly communicated to the result party so that they can understand in which contexts and to what extent the dataset can be reliably used. For example, a synthetic dataset used in a research project must be able to guarantee that discovery is not due to a coincidence of generation but is indeed attributable to a finding that is present in the original identifying data.

Most synthetic data generators offer no formal guarantee with regard to disclosure control. On the contrary, recent developments in this area have shown that these generators are subject to blatant disclosures of identifying data in the same way as any other data product. This limitation can be overcome by combining these generators with an output privacy enhancing technology such as differential privacy.

Any promise regarding the existence of a dataset that reproduces all the properties of the original dataset while also guaranteeing that no identifying data will be disclosed is false and not technically possible.

Promises and maturity

Synthetic datasets make it possible to disseminate microdata, particularly to the public, while limiting the risk of identifying data being disclosed. They offer a flexible and easily implementable solution and open up numerous new opportunities for using the identifying data held by an administration. However, this can only be achieved by combining synthetic data generators with an output privacy enhancing technology. Synthetic data generation is a mature technology with numerous open source implementations available in the form of R or Python packages. While simulated data is perfectly usable in a productive environment within an administration, the second type of synthetic data, reproducing certain statistical properties of the identifying data, requires scientific research to clarify the extent to which the data generated can be reliably used. This uncertainty means there is still debate over whether they can be used by an administration.

4 Use case 1: Platform for the confidential use of data science without data transfer

4.1 Context and objectives

Federal authorities possess vast quantities of data that is essential for the performance of their public interest mandate. The Federal Administration invests considerable amounts of time and resources in collecting, selecting and consolidating these datasets, which constitute a reliable and unique source of information. Despite their intrinsic value, however, they are still not fully exploited for the benefit of the public good. When it relates to a natural person or legal entity, this data is generally restricted to the use for which it was originally collected, in line with the rules governing data protection and statistical secrecy. However, secondary use is possible for purposes that are not related to specific persons, in particular for research, planning or statistics ([Art. 39 para. 1 FADP, SR 235.1](#)). Any such request is examined in detail to evaluate the legitimacy of the request and the associated risk of disclosure. As a result, the administrative process required to obtain access to this data in such cases is long and complex.

The full potential of the data collected by the Federal Administration can therefore only be realised if it is subsequently reused in a range of different fields:

- Firstly, other federal offices and administrative units can use this data to provide effective support for public policy development based on objective indicators. For example, the Federal Statistical Office (FSO) can supply demographic and/or death/mortality data at the individual level to enable the Federal Office of Public Health to adapt its decisions proactively and based as closely as possible on the actual situation. In education, the data on student pathways held by the FSO can be of interest to the cantons, giving them a better understanding of training needs at national level.
- Secondly, research laboratories can use this data to advance knowledge and boost Switzerland's competitiveness through innovation. For example, university researchers could analyse the data held by federal offices to develop innovative solutions to combat poverty or reduce school dropout rates.
- Thirdly, the private sector can use this data to support the health of Switzerland's economy. For example, Swiss companies could use the data held by federal offices to enrich their own data and make strategic decisions regarding investment and product development.
- Finally, this data can also be shared with the governments of other countries to improve the response to global crises such as COVID-19. For example, Switzerland could share epidemiological data with other countries to help coordinate an effective global pandemic response.

However, this secondary use of the data must be politically and socially acceptable, ensuring that the advantages it brings do not come at the expense of Switzerland's citizens, households and companies, as there could otherwise be a significant reduction in people's trust in federal statistical bodies and/or their willingness to participate in surveys. To retain trust in government authorities, it is thus crucial to limit the risk of identifying data being disclosed while they are being used either within or outside the Federal Administration.

This use case explores the scope for reusing collected data while also controlling the risk of identifying data being disclosed. It is important to note that this service requires the data to be collected by a trusted central service. In the event of a potential secondary use of this data, the use of privacy enhancing technologies helps to maintain or even increase trust among people who have entrusted their data to this service. Ensuring trust in this central service is an interdisciplinary governance issue involving the implementation of technical, legal and organisational measures. The coordination of these measures is addressed in section 6.3 of this report.

This project takes the form of a platform that authorised parties who wish to use data for secondary purposes can connect to remotely in order to create data products with a controlled risk of disclosure, and without the user seeing the original data at any stage. The identifying data therefore does not flow between the data owner and the data user. The ultimate aim of this platform is either to facilitate the secondary use of identifying data in cases where this is deemed too risky or the sharing of such data is prohibited by law, or to enable the rapid development of data analysis algorithms in anticipation of a formal data sharing process. The use of technologies to enhance output privacy, such as differential privacy, enables the risk of disclosure to be precisely controlled irrespective of how the identifying data is used. The platform therefore makes it possible to:

- retain public trust in government authorities by limiting the risk of adverse consequences for individuals during the secondary use of identifying data;
- grant much faster access to protected data for purposes other than those for which it was collected. In this case, the platform can for example serve as a transitional solution while waiting for a research laboratory to be given access to the original data. To maintain privacy, algorithm results are perturbed using differential privacy techniques. Extensive perturbation renders the risk of disclosure marginal, so access to the platform can be granted on the basis of minimal legal regulation. The results obtained are therefore subject to a high degree of uncertainty, meaning that their usefulness is limited. This still represents a valuable opportunity, however, as it enables researchers to start work quickly by testing their code and experimental protocol.

The platform is similar in principle to a remote access service, but differs from traditional solutions firstly by ensuring that the identifying data are never disclosed to the user, and then due to the lack of any need for human verification of the algorithms executed by the user. With regard to this last point, remote access platforms require systematic monitoring of the results by experts in disclosure risk, which is both costly and slow. The use of differential privacy therefore makes it possible to automate the entire process while also ensuring that the risk of disclosure is controlled thanks to a mathematical quantification methodology that is future-proof and resistant to technological developments.

To emphasise the potential of this use case for the Federal Administration, here are a few concrete examples of applications within the Federal Department of Home Affairs:

- **Death/mortality statistics:** The usefulness of the data collected by the FSO for death statistics extends beyond statistical purposes, as it also enables researchers to assess the impact of certain factors such as the COVID-19 crisis on mortality. For example, the University of Milan conducted a study of the Swiss population based on a delivery of this data, which is both personal and sensitive. To enable this, a delivery contract had to be drafted and signed transferring the responsibility for protecting the data to the university requesting it: this measure does not limit the reputational risk for the Federal Administration in the event of a data leak by the university. Using the platform described in this use case would make it possible to control this risk, while also making it significantly easier for researchers to conduct this type of study by easing the administrative burden.
- **Swiss Earnings Structure Survey:** The data collected for this survey is of huge public interest, as evidenced by the extensive media coverage of this topic. Against this backdrop, to supplement its conventional statistical production the FSO developed Salarium, an online application that allows users to obtain detailed figures based on criteria such as region, gender, sector and weekly number of working hours. Some uses of the information shared on this page, which is the most visited on the FSO website, include individual salary negotiations, negotiations between social partners and monitoring of benchmark salaries. This data is invaluable for Swiss society, which benefits directly from its secondary use, e.g. for research purposes. This data would make it possible to evaluate and adapt political, social and economic measures to ensure they are both effective and fair. The use case presented here would facilitate and accelerate the realisation of

this potential while still respecting the privacy of members of the public, and thus contribute to the acceptance of this secondary use.

- **Digisanté:** This programme aims to promote the digital transformation of the health system. Package 4 of the programme is focused on two objectives: enhancing planning by improving the use of the data held by public health services, and enabling the relevant authorities and stakeholders to implement a pilot of the health system. The second objective aims to make it easier for researchers from the academic community and the private sector to access this data in order to treat health problems. The use case described in this section offers an obvious technical solution to this challenge, making it possible to ensure privacy during the secondary use of health data and thus realise the implementation of a privacy-preserving 'data space for health-related research'.

In summary, the platform presented in this section would boost the potential for the automated use of the data held by the Federal Administration while also guaranteeing the privacy of identifying data, thus limiting both the reputational risk for the Federal Administration and the risk of adverse consequences for the public.

4.2 Principle

Parties and trust schema

The platform architecture is designed to maximise the use of the specific characteristics of trust relationships within a public administration:

- The privacy units transfer their data to the designated input party, generally a federal office, within the framework of a trust relationship established by an existing legal basis.
- Each federal office can call on a trusted computing party, such as the Federal Office of Information Technology, Systems and Telecommunication (FOITT), eliminating the need to use input privacy enhancing technologies. This party is responsible for data security, and must therefore implement the measures required to prevent unauthorised access to the data and the platform. It is also responsible for identifying and authenticating the platform's users.
- The result parties for the data may range from employees of other federal offices to the general public. In all cases, it is assumed that there is no trust relationship between the input party and the result party. Consequently, output privacy enhancing technology, in particular differential privacy, must be used to mitigate the risk of identifying data being disclosed.

This trust model is common in public administrations, which can rely on computing capacity managed directly by government authorities.

Description of the service

Figure 4 shows how the service operates, highlighting the practical scenario in which the data is made available to an authorised user. From the user's perspective, the service operates as follows:

- 1) The user connects to the platform remotely, thereby accessing a virtual environment conceived for the execution of the Python or R programming languages. Expert knowledge of these languages is essential for using the platform, given their prevalence in data analysis in both the private sector and the academic community.
- 2) With the aid of a library, the user can consult an exhaustive catalogue of the datasets available to them, as well as their type and metadata, while also taking note of the privacy-loss budget assigned to them.

- 3) Once a dataset has been selected, the user can download to their computing environment a dummy dataset that is entirely simulated but faithfully reproduces the metadata of the original dataset. This dummy dataset can be used to design and test data analysis algorithms.
- 4) The user can then transmit their completed algorithm to a remote server operated by the Federal Administration. Before being executed, the algorithm is analysed to verify the presence of a privacy control mechanism. If this mechanism is not present, the algorithm is not executed. The server then verifies the user's access rights in respect of the dataset and the adequacy of their privacy-loss budget.
- 5) If the access and privacy conditions are fulfilled, the secure server retrieves the identifying data from a secure storage space.
- 6) The algorithm is then applied independently to the original dataset, and the result is computed while preserving privacy with the aid of differential privacy.
- 7) The protected result is recorded and transmitted to the user. This recording stage enables the user to consult their previous results, thus avoiding repeated use of their budget for the same analysis. Archiving the algorithms and associated results also enables the authorities to monitor users' activity and how the different datasets are used.

This approach enables the user to execute algorithms without having direct access to identifying data, while also giving assurance to the office supplying the data that the risk of identifying data being disclosed is limited.

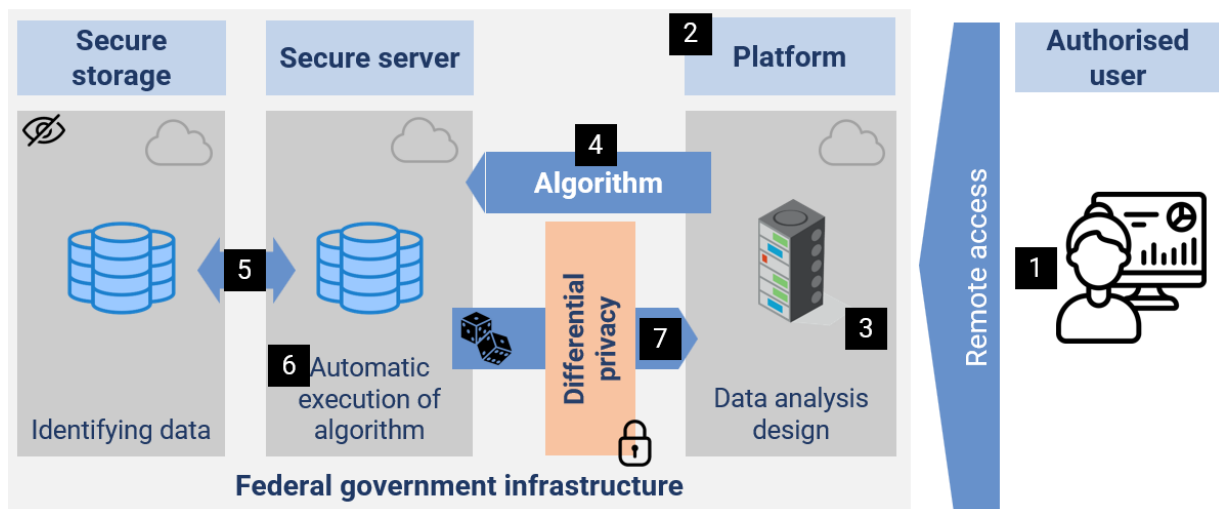


Figure 4: Schematic of a platform that permits the confidential application of data science without data being transferred.

For a data analysis to be authorised, it must include a disclosure control mechanism such as a randomisation algorithm based on differential privacy. The platform does not develop these algorithms internally but relies on recognised open source libraries, such as OpenDP and SmartnoiseSQL for manipulating and transforming data, and DiffPrivLib for machine learning and artificial intelligence, to name just a few.

The secure server is linked directly to the storage space where the datasets are stored. With the democratisation of the 'once-only' principle, using the platform benefits initiatives around the creation of catalogues that classify the datasets held by the Federal Administration and above all catalogue their metadata. For example, the I14Y standard simplifies the generation of dummy datasets for evaluating data analysis algorithms designed by users. Directly integrating these services, as permitted by the 'fso-

metadata' package developed by the FSO, makes it possible to take full advantage of these initiatives by combining them and thus fully exploit the value of the data held by the Federal Administration.

To prove the feasibility of such a platform, the FSO's Data Science Competence Center (DSCC) has developed an **open source** implementation that is available on its GitHub.²⁰ **This platform, developed to ensure its portability on a standard IT infrastructure, is therefore available to all entities of the Federal Administration wishing to test this type of service or even deploy it in a productive environment.** It is accompanied by a report covering all of the technical details (<https://arxiv.org/abs/2406.17087>).

4.3 Limits and challenges

This use case takes advantage of the specific characteristics of trust relationships within the Federal Administration. By working with a trusted computing party, the focus is placed exclusively on guaranteeing the confidentiality of the results produced. This protection is implemented through the use of differential privacy. The main challenge associated with differential privacy lies in determining the privacy-loss budget, a concept used to precisely regulate the trade-off between the usefulness of data and the risk of disclosure by defining a maximum privacy-loss threshold for each user. The choice of this threshold is a question that includes a political dimension and can be formulated as follows: "What is the maximum level of risk to which the privacy units for the data set can acceptably be exposed?", or alternatively "What is the minimum level of usefulness required to guarantee the reliability of the dataset?". These two questions correspond to the principles of 'confidentiality first' and 'utility first' respectively, and thus enable an organisation to define its priorities and set its privacy-loss budget in a transparent and justifiable manner.

With this in mind, we will now look at the example of how data is used in research. The current process is characterised by the slow and cumbersome procedure for gaining access to identifying data, which requires a formal legal framework transferring responsibility for protecting the data to the research laboratory that is using it. In addition to the high administrative cost, this process also acts as a significant brake on the speed of research, a situation that is sometimes incompatible with urgent circumstances such as the COVID-19 pandemic. To illustrate this challenge, we present two particularly relevant use cases for the platform:

- "Facilitation and/or preliminary work": This approach aims to supply data to a researcher quickly, enabling them to start their work without delay. This gives them the opportunity to verify the feasibility of their analyses and, if they are validated, to prepare and test their algorithms while awaiting completion of the administrative process for obtaining the original data. The advantages are twofold:
 - The user can accelerate their research project.
 - The Federal Administration can save on resources if the analyses prove not to be feasible.
- Make possible: This approach relates to the provision of data that was previously deemed too sensitive, but which can now be used since it is never disclosed in its original form and the risk of disclosure is controlled. The benefits are as follows:
 - The user can analyse data that is not normally accessible.

²⁰ <https://github.com/dscc-admin-ch/lomas>

- The Federal Administration can exploit the full potential of the data collected. There are also positive reputational implications, as the benefits offered by the financial investment in collecting and consolidating the datasets extend beyond the organisation responsible.

These two use cases require a small privacy-loss budget to offset the risk, either due to the sensitivity of the data or due to the need for short timescales, which makes it difficult to implement extensive controls on potential users or conclude formal contracts.

Finally, while the risk of disclosure is evaluated and controlled for each user based on the privacy-loss budget, no technical measure can prevent the possibility of several users joining forces and combining their individual budgets with the aim of extracting identifying data from the platform. Any such collusion must be explicitly prohibited by legal regulations governing the use of the platform. Consequently, it is not possible to completely avoid having a legal framework to govern the use of the platform. As long as users' privacy-loss budgets remain limited, this framework is probably less restrictive than the agreements required for sharing identifying data. This translates to resource savings and much easier access for users. Within the Federal Administration, identity control and user traceability can be assured by means of authentication protocols such as eIAM.

5 Use case 2: Secure data enrichment based on private linkage

5.1 Context and objectives

The systematic collection of identifying data and its limits

The systematic collection of data now extends to all areas of society, reflecting a significant evolution in the way that information is gathered and used. In the private sector, this practice is particularly evident within banking institutions, which meticulously collect data on financial transactions and their clients' consumption habits. Providers of smart devices also conduct intensive data collection, notably in areas such as personalised health, with connected devices monitoring and recording individuals' physiological parameters in real time. This trend also extends to sectors such as energy efficiency, with intelligent sensors measuring and analysing energy consumption in residential and commercial buildings, and mobile telephony, with operators gathering data on individuals' mobility.

Systematic data collection is also conducted in the public sector, notably in the field of medicine, albeit within a strict legal framework that is often based on obtaining explicit consent (Art. 7 para. 1 of the Human Research Act (HRA)). For example, it is becoming increasingly common for patients being admitted to hospital to be asked to consent to their data being used in clinical or epidemiological studies.

Non-governmental organisations (NGOs) also play a major role in the collection of data for humanitarian and development purposes. They gather information on migratory flows in order to gain a better understanding of population movements and identify needs with regard to assistance and protection for migrants. They also measure economic and social development in developing countries in order to evaluate the impact of aid programmes and help shape development policy. Finally, they participate in the management of international crises such as epidemics or natural disasters by collecting the data required to evaluate humanitarian needs and coordinate emergency responses.

It is important to note, however, that the systematic nature of this collection inevitably leads to bias, duplications and, on a more general level, inaccuracies. Consequently, their use in sensitive areas, over and above simple targeted advertising, for example, is a delicate undertaking that requires in-depth analysis and modelling along with measures to guarantee security and privacy.

At the same time, the Federal Administration is making substantial investments, in terms of both time and resources, in the collection, selection and consolidation of datasets that are classed as reliable and trustworthy. This data represents a unique source, comparable only to certain clinical studies in the health field. However, certain information that is crucial for the functioning of the Federal Administration remains unavailable or uncollected due to the lack of a legal basis or because it is not accepted by a society that considers the associated risk for individuals to be too high. The full potential of the data held both by different bodies of the Federal Administration and by researchers, NGOs and the private sector can be realised by enriching it with the data held by the public sector. This enrichment process is achieved through linkage, which involves combining two (or more) data sources by linking the information relating to the same identifiable person. This makes it possible to analyse the results of this linkage or use them to train machine learning and artificial intelligence models. However, this practice must at all times preserve the privacy of the identifiable units whose data is being used.

Data linkage: opportunities for the public sector

Data linkage offers significant advantages for the public sector. First of all, it permits the enrichment of data held by other administrations, thus facilitating linkage between different administrative entities such as federal offices or other administrative levels such as cantons and communes. For example, bodies such as the Federal Office of Public Health (FOPH) or the Federal Office for the Environment (FOEN) are able to evaluate the impact of public health or environmental policies on mortality using data collected by the Federal Statistical Office (FSO). Enrichment with data held outside the Federal Administration

also offers additional perspectives. For example, data from external stakeholders such as the banking sector would allow economic indicators of consumption to be estimated more accurately. The data collected by energy suppliers could also improve the management of energy efficiency development programmes or help draw up effective rules in the event of shortages. Finally, the mobility data held by mobile phone operators can, for example, be used to optimise modes of transport and transport infrastructure, and could be crucial for managing national crises such as epidemics or natural disasters.

Data linkage also offers significant advantages for stakeholders outside the public sector. In the research field, particularly in the area of health, numerous stakeholders hold important datasets that may be biased due to poor collection campaigns. The deficiencies in these datasets can be rectified by enriching them with reliable data such as that held by the FSO, thereby re-establishing the scientific credibility of the results of studies conducted using these imperfect datasets. This approach represents a convincing solution for partially resolving the problem of reproducibility facing the contemporary scientific community. The FSO is currently permitted to use linkage for planning, statistical and research purposes (Art. 14a FStatA), but this is extremely costly in terms of both administrative and technical resources. The use of privacy enhancing technologies presents an opportunity to offer this service on a large scale while limiting the financial burden that it places on the Federal Administration.

As for NGOs and intergovernmental collaboration, international institutions are also seeking to promote the public good beyond national borders. However, exchanging data at an international level or with certain intergovernmental organisations is often hampered, or made uniquely complex, by the existence of strict regulations. The identifying data held by one or more countries could be used in areas such as international commerce or migration to tackle global challenges like climate change.

Furthermore, any private-sector company that systematically collects identifying data could enrich this data and make it more reliable by linking it with data held by the Federal Administration. This would necessitate a clear legal framework and would provide a boost to the Swiss economy, as it would stimulate innovation by validating the data products created.

This type of linkage, primarily involving sensitive data, must be both politically and socially acceptable, while also guaranteeing that the advantages arising from its use do not come at the expense of Switzerland's citizens, households and companies. To retain trust in government authorities, measures must be taken to limit the risk of identifying data being disclosed during the linkage process, either within or outside the Federal Administration, and to guarantee that it is impossible to draw reliable conclusions about individual privacy units.

A secure data enrichment service based on private linkage

This practical case examines the possibility of linking data from sources who do not have the right to, or simply do not wish to, share identifying data with each other, but who are nevertheless seeking to create a product that combines the two datasets. This scenario takes the form of a service that the authorised parties who are interested in data linkage can connect to remotely in order to make their data available, enrich it through linkage, then create data products for which the risk of disclosure is controlled. The entire process is conducted without the contributors being able to access the other parties' original identifying data at any time. The main aim of this service is either to enable the Federal Administration to enrich its own datasets with the aid of external data sources, or to make available the datasets held by the Federal Administration in order to increase the reliability of data from external sources.

In other words, this service facilitates collaboration between two input parties who do not trust each other, with the aim of pooling their data and creating a data product that uses both sources. The absence of a trust relationship between these parties is key, as it makes it necessary to ensure that the technical solution does not allow the different stakeholders to access the linkage results directly. The identifying data is never made available in its original form, and only the output from the data product created from the linkage results is transmitted to the result party after the application of measures to control the risk

of disclosure. Consequently, an output privacy enhancing technology, in particular differential privacy, must be combined with an input privacy enhancing technology.

Input privacy guarantees that the parties involved never have direct access to the original identifying data used for linkage, and that the code executed by the service is indeed that agreed by the two input parties. At the same time, output privacy makes it possible to precisely control the risk of disclosure associated with the creation of the data product using the enriched data. This ensures that no reliable conclusions can be drawn about individuals based on the linkage results. Conclusions can only be drawn about groups of people, but their benefit for the public good must still be evaluated by determining who uses these results and for what purpose. It is worth noting that output privacy enhancing technology renders any attempt at individual profiling, targeting or simply extracting information about specific individuals unreliable. The service architecture described in the next section has been selected to allow for an optimal combination of these two types of privacy enhancing technologies.

Due to its role and specific architecture, the service performs several key functions:

- It helps retain public confidence in government authorities by guaranteeing that users do not have access to original data, thus limiting the risk of adverse consequences for individuals resulting from the use of identifying data.
- It helps government authorities fulfil their mandate more effectively by enabling them to access new data sources.
- It benefits society as a whole by making available the reliable data sources held by the Federal Administration.

In summary, the service makes it easier to pool the data held by the different stakeholders in Swiss society, enabling each of them to benefit from these resources while also guaranteeing the privacy of individuals' identifying data and precisely controlling the risk of adverse consequences for the public.

5.2 Principle

Parties and trust schema

The service architecture is defined by the absence of trust relationships between the different contributors and users, as illustrated in Figure 5.

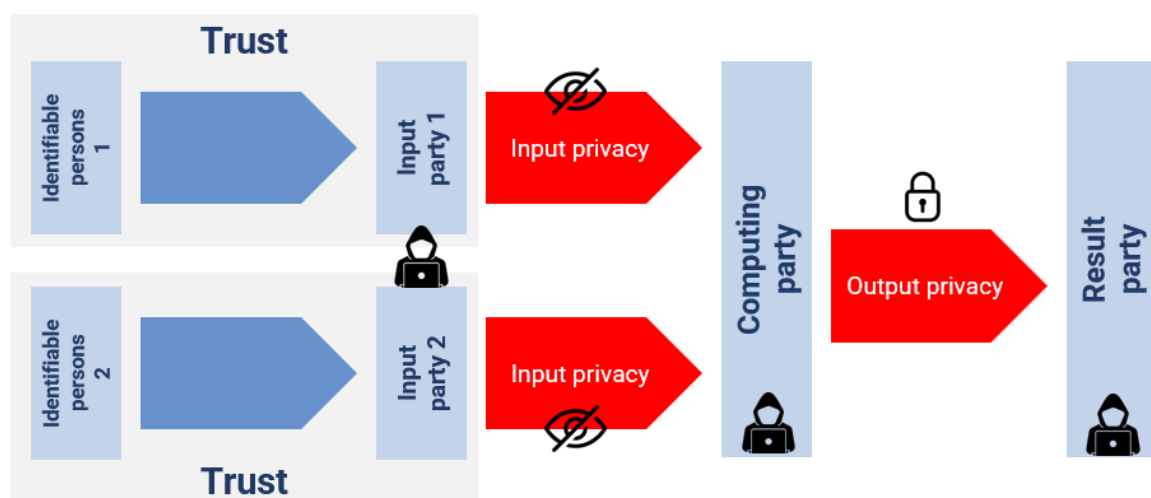


Figure 5: Presumed trust schema for the use case "Secure data enrichment based on private linkage".

The parties involved are broken down as follows:

- 1) Two or more groups of identifiable persons whose data is collected by two or more independent input parties. These collections are separate, do not cover the same information and include a subset of common identifiable individuals. One of the input parties wishes to combine the information from the two collections to create a data product that is enriched by the two sources.
- 2) Input party: One of the input parties is a body of the Federal Administration, whereas the second may belong to any stakeholder category in Swiss or international society. To guarantee the privacy of the identifying data held by the two parties, it is essential to prescribe that neither input party may access the other's identifying data without the application of disclosure control technology.
- 3) Computing party: This party cannot be deemed trustworthy by the two input parties. Traditionally, the FOITT or any government IT service provider within an administration is deemed trustworthy thanks to the applicable legal regulations and IT security measures. In a multiparty context, however, at least one of the two input parties does not have a trust relationship with the Federal Administration, meaning that it must not be technically possible for the computing party to access the input parties' data. This absence of trust does not necessarily mean that the government computing infrastructure must be used. Computation can also be outsourced to public cloud service providers. Consequently, the technical solution for input privacy must offer adequate security and privacy guarantees for this scenario.
- 4) The result party: In this case, the result party for the data product is assumed to be one or both of the input parties. Given that there is no trust relationship between these two parties, it is therefore necessary to protect the product with the aid of a privacy enhancing technology that ensures **input privacy**.

Choice of privacy enhancing technologies

The multiparty nature of this use case restricts the choice of input privacy technologies to the following three options: trusted execution environments, secure multiparty computation and homomorphic encryption. The choice of technology is guided by several criteria, in particular compatibility with the chosen output privacy technology (differential privacy), the availability of open source solutions to limit the risk of 'vendor lock-in', and how easy a service is to implement and deploy. A detailed analysis of the advantages and drawbacks of each technology is provided in section 3. The motivation behind the choices in this specific use case are described in brief below:

- Secure multiparty computation does not yet have sufficiently mature open source implementations to be used in this type of application. Combining differential privacy with secure multiparty computation protocols also still presents a number of challenges, in particular distributed random number generation. While some technologies do exist, they are still at the research stage.
- There are open source multiparty solutions for homomorphic encryption, but combining these with differential privacy in the absence of a trusted third party requires a randomisation mechanism to be applied directly to the encrypted data in order to decrypt the data product and apply differential privacy. This is a complex task for which no mature solution is currently available.
- Trusted execution environments currently make it possible to deploy services as described in use case 1 with a limited number of modifications. This use case can therefore be deployed directly within an enclave, offering a simple solution that combines technologies to enhance both

input privacy and output privacy. In addition, public-private initiatives such as Gramine²¹ offer open source solutions that can be ported to different cloud service providers. In combination with a service for attesting the integrity of the executed code, trusted execution environments represent the most complete solution for the implementation of this use case. Unlike the other two technologies, however, these environments do not provide any mathematical guarantee of security and necessitate trust in the hardware manufacturer, for example Intel or AMD.

For this use case, execution environments appear to be a natural solution that offers considerable synergies with the use case presented in the previous section, whose components can be reused directly to create data products once linkage has been carried out.

Description of the service

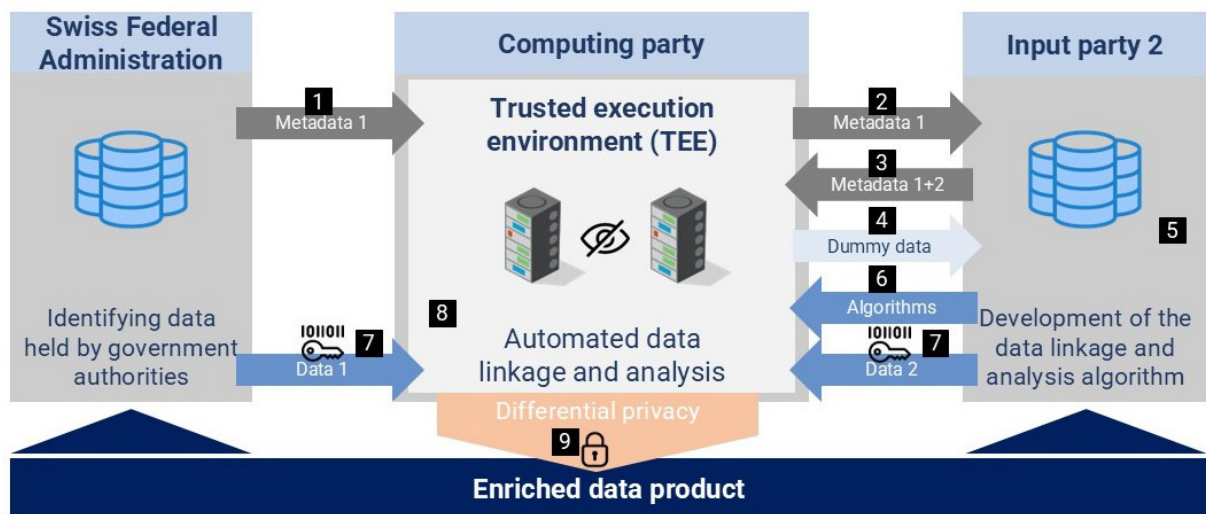


Figure 6: Schematic of a service that enables data products to be created from data enriched through private linkage while also controlling the risk of disclosure

Figure 6 shows how the service operates. By default, we work on the principle that input party 1, represented in Figure 6 by the Federal Administration, makes its datasets available, while input party 2 requests that a data product be created by also contributing its own data. We presume that input party 1 and input party 2 have agreed on an acceptable privacy-loss budget for the creation of this product and the sharing of their databases in advance. Finally, it is assumed that the Federal Administration is responsible for operating the service and must therefore establish the necessary authorisations, allocate the privacy-loss budget and deploy the enclave.

From the perspective of input party 2, the service operates according to the following stages:

- 1) Input party 1 uploads its metadata to the trusted execution environment.
- 2) Input party 2 retrieves the metadata from the dataset supplied by input party 1.
- 3) Input party 2 expands the catalogue of metadata by integrating its own metadata, and standardises the encoding of its data by aligning it with that of input party 1. Finally, input party 2 uploads the consolidated catalogue of metadata to the trusted execution environment.

²¹ <https://gramineproject.io/>

- 4) The consolidated catalogue of metadata is used to generate a dummy dataset that reproduces the characteristics of datasets 1 and 2. This dummy dataset is transmitted to input party 2 for the development and local testing of the linkage and product creation algorithm.
- 5) The data product creation algorithm and comparison rules between datasets 1 and 2 are developed using the dummy dataset.
- 6) The data linkage and analysis algorithms are uploaded to the trusted execution environment. The algorithms are tested in this secure environment using the dummy dataset.
- 7) The two sets of identifying data held by each of the input parties are uploaded to the trusted execution environment. This transfer is encrypted to prevent any unauthorised access.
- 8) Checks are carried out to verify the presence of a differential privacy mechanism and the existence of an adequate privacy-loss budget for the execution of the data product creation algorithm.
- 9) The linkage and data product creation algorithms are executed automatically. The product is then sent to the input parties in encrypted form.

Stages 7, 8 and 9 are similar to those described in use case 1 and are based on the deployment of the use case described in section 1 in a trusted execution environment. This process therefore enables input party 2 to create its data product from the linkage results without ever accessing identifying data, while also providing input party 1 with a guarantee that the data product created presents only a limited risk of identifiable information being disclosed.

5.3 Limits and challenges

The use case presented here is fundamentally based on the harmonisation and standardisation of the metadata supplied by the two input parties. These are assumed to be shared between the two input parties, meaning that column names, content descriptions and all possible values need to be disclosed to each party. It is essential to note that neither linkage nor the automated use of personal data can be realised without this preliminary stage. To fully exploit the potential of the data held by the Federal Administration, it is vital that the effort required for this harmonisation is realised, and we wish to stress once again the importance of initiatives such as I14Y if we are to take full advantage of the potential of the data held by government authorities. Privacy enhancing technologies (PETs) are emerging as a technical solution permitting the secondary use of the data held by the Federal Administration in a way that preserves privacy once standardisation has been achieved. These two elements, namely PETs and the harmonisation of metadata, are fundamental to modern data science and thus inextricably linked, and must be implemented jointly in order to make full use of the data held by Swiss society.

This second use case is directly based on the tools presented in use case 1. Consequently, it is exposed to the same difficulties and challenges, notably regarding the choice of the privacy-loss budget and the implementation of the associated legal framework. Over and above the legal and methodological challenges, however, this use case relies on the use of specialised IT hardware, namely computing chips that permit the strict separation of users (TEE/confidential computing). These types of machines are now part of the product offering of most public cloud service providers such as Azure, AWS, Google Cloud, etc. However, they are not yet offered by the Federal Administration's IT services. Investment in infrastructure and human resources would be required to implement and offer this type of service within the Federal Administration. However, the secure nature of these services could justify a decision to make direct use of the above external providers rather than making this kind of internal investment. The choice between developing internal capacities or relying on public cloud services is therefore a political question related to the governance and use of data. To fully exploit the potential of the data held by government authorities, therefore, the Federal Administration must either invest in the infrastructure required to make this data available using the technologies described above, or accept a loss of governance and establish the legal basis required to use public cloud services such as trusted execution environments. A new legal basis is also required in order to use a service such as that described in this section.

6 Legal and governance framework

This section illustrates how privacy enhancing technologies (PETs) fit into the existing legal framework, and the extent to which they may offer potential approaches for facilitating the secondary use of identifying data while also preserving the privacy of individuals. This review of the legal framework is thus in no way intended to be exhaustive. In the event of the concrete implementation of the use cases described in this report, all provisions relating to data protection must be applied, specifically the Data Protection Ordinance (SR 235.11, DPO) and the confidentiality requirements laid down in international instruments governing administrative assistance in tax matters.

6.1 Processing of identifying data – review of the legal framework

The existing legal framework governing the processing of identifying data (see section 2.1) makes a distinction between the personal data of natural persons and data relating to legal entities, and only addresses any grouping of the latter indirectly. The legislation ensures the protection of a group by guaranteeing the protection of each individual within this group. From a technical point of view, however, protecting a group requires specific and more restrictive measures than those used to ensure the protection of each individual within the group. With this in mind, this report considers identifying data in a general sense, in other words information associated with any type of identifiable unit (see section 2.1 for precise definitions).

It is worth recalling here that the concept of **processing** in the sense of the law (Art. 5 let. d FADP) covers the entire data lifecycle and the entire privacy preserving data science process,²² in other words collection, storage, retention, use, modification, disclosure, archiving and finally deletion or destruction. This processing by the Federal Administration is based in particular on two key principles:

- **Anonymisation of identifying data:** The existing legal framework is based on the principle of anonymisation and deletion, to guarantee the privacy of identifying data. Data considered to be anonymised falls outside the scope of data protection legislation.
- **Primary use of data:** The existing legal framework restricts the use of data collected by the Federal Administration to that provided for by law. All data processing requires an explicit legal basis and a clearly defined purpose. Any secondary use for other purposes is in principle impossible, with the notable exceptions of statistical, planning or research purposes or if a legal basis is created to authorise the processing of previously collected data for another purpose.

Anonymisation and the prohibition of secondary use reflect the principles of the **rule of law** (Art. 5 of the Swiss Constitution (Cst); SR 101), which stipulates that all activities of the state must be based on law, conducted in the public interest and proportionate to the ends sought. These measures also ensure the **protection of privacy**, in other words that “every person has the right to be protected against the misuse of their personal data” (Art. 13 Cst), and the upholding of **fundamental rights** by the state (Art. 35 Cst). Consequently, the legal framework for data processing is shaped by the principles of **good faith, legality, proportionality, purpose limitation and equality**.

With regard to the processing of identifying data relating to natural persons, in other words **personal data**, the Federal Act on Data Protection (FADP, SR 235.1) applies to both private-sector entities and federal bodies. The law implements the principles of anonymisation and purpose limitation in the following way:

²² <https://www.bfs.admin.ch/bfs/en/home/dscc/dscc.html>

- Only the processing of **non-anonymised** personal data, in other words any information relating to an identified or identifiable natural person (Art. 5 FADP) is governed by the FADP. Any anonymised data therefore falls outside the scope of this law.
- The FADP explicitly requires the existence of a **legal basis** for the processing of this personal data by a federal body (Art. 34 FADP). The requirement for a legal basis in a formal or substantive sense depends on the degree of sensitivity of the data and the purpose of the processing. If the data processing may lead to a serious violation of the data subject's fundamental rights, a formal law is necessary, whereas if the purpose of processing is the performance of a task already defined in a formal law, substantive law may suffice. The FADP also covers cases in which a federal body processes personal data jointly with other federal bodies, with cantonal bodies or with private persons, as described in use case 2 of this report, stipulating that in such cases the Federal Council shall regulate the control procedures and responsibilities (Art. 33 FADP).

The processing of **legal entities' data** by a federal body is covered by a transitional provision of the new FADP (Art. 71 FADP). In principle, however, the Government and Administration Organisation Act (GAOA, SR 172.010) must be considered. These laws implement the principles of anonymisation and primary use in the following way:

- By analogy with Article 5 FADP, legal entities' data that is considered to be anonymised falls outside the scope of the law.
- Article 57r of the GAOA confirms the requirement for a defined purpose for the processing of legal entities' data, including sensitive data, by federal bodies. The purpose of processing must correspond to the tasks of the federal body as defined in a formal act of parliament.

In terms of the formal legal basis for specific federal bodies, one especially relevant example in the context of this report relates to **federal statistics**:

- The Federal Statistics Act (Art. 18 para. 3 FStatA) requires the **anonymisation** of data products, which must no longer contain any elements that would identify the individuals, households, companies or establishments surveyed. Any violation of this principle may potentially give rise to a liability for the Federal Administration under the Federal Act on the Liability of the Federal Government, Members of its Authorities and its Public Officials (SR 170.32), and if the legal conditions are met could result in criminal sanctions and direct repercussions for the reputation and image of the Federal Administration.
- Federal statistical bodies are governed by the **Federal Statistics Act** (FStatA, SR 431.01) and the Ordinance on the Conduct of Federal Statistical Surveys (SR 431.012.1). The data collected by federal statistical bodies is primarily used for statistical purposes, but there is a legal basis for its secondary use in the form of data delivery or linkage exclusively for statistical, planning or research purposes. The application of this legal basis requires a contract to be signed that generally transfers responsibility for data protection to the recipient of the delivery. The new draft ordinance on federal statistics promotes the multiple use of data through the implementation of the 'once-only' principle, but its use is largely limited to statistical, planning and research purposes.

Finally, in the area of health, the **Human Research Act** (HRA, SR 810.30) protects data that is not considered to be anonymised by implementing the principles of anonymisation and primary use in a similar way.

6.2 Limits of the existing legal framework

6.2.1 Current status of attacks aimed at extracting identifying data: anonymisation no longer effective

The relevance of the legal framework described above has to be viewed in the context of recent developments in techniques to extract identifying data from a data product. What was once seen as a purely theoretical risk, namely an attack on privacy triggered by the disclosure of attributes such as membership to a sensitive population group or the direct re-identification of an identifiable person, has now become a reality. Furthermore, the technical implementation of such attacks is no longer an abstract vision but is now entirely possible on a large scale, giving rise to three categories of attacks on privacy.

Firstly, linkage attacks involve combining a dataset with one or more sources of auxiliary information, the aim being to link information relating to the same identifiable unit with a view to their re-identification. This microdata may either be obtained illegally by exploiting a security flaw, or via legal methods such as a delivery of pseudonymised data, reconstruction or the extraction of data from a publication. As such, this type of attack primarily concerns what is commonly referred to as microdata, in other words data that is directly linked to identifiable persons. With the democratisation of the systematic collection of this type of data, any protection tool needing to anticipate the nature and scope of the auxiliary information available to an attacker is doomed to fail. The re-identification of users' film viewing history on the Netflix platform by researchers at the University of Texas at Austin²³ who were taking part in a data science competition shows just how easy it is to carry out this type of attack and underlines the impossibility of anticipating how attackers use the auxiliary data that is available to them.

Linkage attacks almost exclusively involve microdata. Data products involving a form of aggregation or generalisation are subject to another type of attack known as a reconstruction attack. The aim of such attacks is to extract the microdata used to create a data product. This type of attack was previously perceived as a potential risk but is now a real vulnerability. This is demonstrated by the experience of the U.S. Census Bureau,²⁴ which was able to successfully reconstruct and re-identify a significant proportion of the US population using the 2010 census tables, despite the use of conventional disclosure control techniques such as aggregation and the creation of thresholds that were intended to guarantee a minimum number of individuals in each group.

The third and final type of attack, known as a membership attack, aims not to re-identify an individual but simply to determine whether they belong to a potentially sensitive group within the dataset. These attacks are particularly attractive for attackers, as the conditions required for them to succeed are a lot simpler to put in place than for the two other types. The associated risk is particularly high in the area of public health, where simply belonging to a group of individuals suffering from a rare illness or condition may disclose extremely sensitive information.

These three categories of attack expose the vulnerability of all data products by circumventing the methods used to control the risk of disclosure. No technical solution can offer an absolute guarantee that 100% of these attacks will fail. This therefore means that it is not technically possible to guarantee anonymisation as described in the FADP and the GDPR, since this necessitates the irreversible modification of the identifying data, something that is technically impossible to achieve. However, anonymisation remains possible from a legal perspective if, based on the context in which the processing took place, the risk of these attacks succeeding is deemed to be sufficiently low.

²³ https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf

²⁴ www2.census.gov/library/working-papers/2023/adrm/ces/CES-WP-23-63.pdf

The provision of any data product, be it microdata or any other aggregated form, exposes individuals to a non-negligible risk of identifying data being disclosed. The only way to reduce this residual risk to zero is by not creating and transmitting the data product. This disclosure risk therefore represents the price to be paid in order to benefit from the advantages offered by the data product. It is a trade-off between privacy and utility, emphasising that one can only be achieved at the expense of the other. When creating any data product, the disclosure risk therefore has to be weighed against the benefits (in terms of utility) in order to formally quantify this trade-off, making use of output privacy enhancing technologies such as differential privacy.

6.2.2 Consequences for the legislation governing the use of data held by government authorities

The technical impossibility of guaranteeing that anonymisation is strictly irreversible shows that technical advances have rendered this principle obsolete. This means that the current legal framework is now limiting the scope to realise the full potential of the data held by the Federal Administration. What is required is a more nuanced legal framework based on the explicit quantification of the residual risk of disclosure, something that is now technically possible through the use of privacy enhancing technologies such as differential privacy.

At the same time, the principle of purpose limitation is also massively restricting opportunities to generate value from the data held by the Federal Administration. The current legal framework therefore offers very little room for manoeuvre with regard to the secondary use of this data for data science activities that are compliant with the Federal Administration's code of practice for human-centric and trustworthy data science (point 3 of the Federal Council decision of 2 December 2022; EXE 2022.2746),²⁵ or for its joint processing as described in use case 2 of this report or in the example "Global tables in crisis situations" on page 17 of the report "Data science for the public good – use cases in policy development" (point 2 of the Federal Council decision of 2 December 2022; EXE 2022.2746).²⁶

The advent of privacy enhancing technologies (PETs) means that an expansion of the existing legal framework governing purpose limitation can be considered. PETs offer a technical solution for facilitating joint use without data being shared and with formal controls limiting the risk of disclosure during the secondary use of data. As such, they limit the risk of adverse consequences for individuals that could arise from an attack on the privacy of identifiable units during the secondary use of their data. PETs therefore permit the privacy-preserving secondary use of data by federal bodies for purposes that are not related to specific persons, subject to the provisions set out in Article 39 of the FADP.

However, PETs alone cannot guarantee that this data will not be misused, in other words that it will be processed in a way that respects all fundamental rights. Nor do they guarantee that an intended secondary use of data will be ethical and beneficial for the public good. The creation of this data product must therefore be viewed within a broader context: in addition to privacy protection, its nature, the data required and the recipient must also be considered. A governance framework that encompasses all of these aspects and their associated risks is needed to ensure that any secondary use of data respects the rule of law.

6.3 The Five Safes: A governance framework for privacy preserving data science

The technical impossibility of anonymisation and the restrictions imposed by the principle of purpose limitation on creating value from data mean that a more nuanced framework is now required that takes account of the full spectrum of risks of attacks on fundamental rights and on the privacy of individuals.

²⁵ <https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees.assetdetail.29325685.html>

²⁶ <https://www.news.admin.ch/newsd/message/attachments/87132.pdf>

The binary nature of these principles, which are based on the concept of compliance, fails to capture the complexity of disclosure risk as it is recognised today, and does not allow the Federal Administration to make optimum use of its data to respond to current challenges. It therefore seems necessary to propose a framework for assessing the entire **spectrum** of risks and the technical, organisational and legal measures implemented to control them. Tools need to be presented that offer an **interdisciplinary and global vision** of every intended secondary use of data and allow informed decisions to be made regarding the risk-utility trade-off, especially in respect of the risk of identifying data being disclosed.

Developed more than 20 years ago, the Five Safes governance framework is widely used by national statistical offices and public data archives for research purposes. The governance framework evaluates any intended secondary use of data on the basis of five principles, namely the nature of the intended use, the data involved, the recipient(s), the environment and finally the result. It therefore covers all risks associated with the creation and sharing of data products derived from identifying data. Users include the national statistical offices of the United Kingdom (ONS UK), Australia (Australian Bureau of Statistics) and New Zealand (Stats New Zealand), along with the UK Data Service,²⁷ Research Data Scotland²⁸ and the Data Innovation Program of the province of British Columbia in Canada.²⁹

A tried-and tested approach for regulating the processing of sensitive identifying data, the Five Safes pre-date the emergence of PETs, but thanks to their comprehensive and flexible nature they can also be used for projects that rely on the use of PETs. Figure 7 sets out in detail where this framework could sit within the context of human-centric and trustworthy data science.

²⁷ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>

²⁸ <https://www.researchdata.scot/our-work/data-explainers/what-is-the-five-safes-framework/>

²⁹ <https://www2.gov.bc.ca/gov/content/data/about-data-management/data-innovation-program/privacy-security>

The Five Safes: A governance framework to unlock the potential of government data in a secure and ethical way

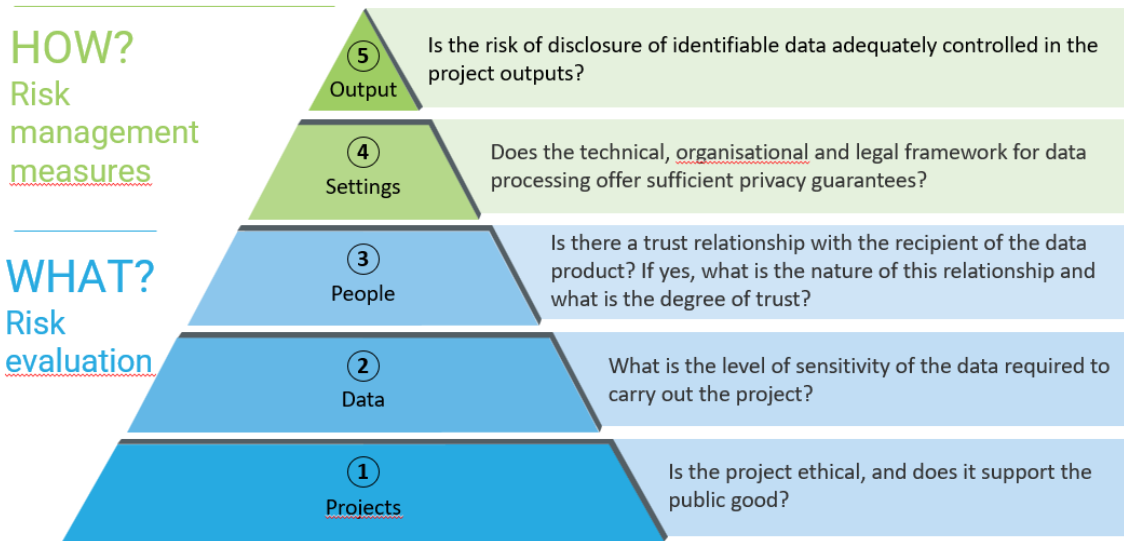


Figure 7: The Five Safes – A governance framework for secondary use of the data held by the Federal Administration that respects fundamental rights and privacy in particular

This governance framework therefore enables us to assess the context for the creation of a data product on the basis of the following principles:

- 1) **Safe projects:** Is the project ethical, and does it support the public good? Is the project compliant with human-centric and trustworthy data science that supports public policy?³⁰ It is therefore necessary to adopt a classification of data products based on their sensitivity and their benefits and risks for society. The European Union conducted a similar exercise for artificial intelligence applications in developing a legal framework for AI.³¹ This type of classification can be implemented to regulate the secondary use of data within the Federal Administration and is comparable to the risk assessment component of a data protection impact assessment (DPIA, Art. 22 FADP).
- 2) **Safe data:** What is the level of sensitivity of the data required to carry out the project? This relates to the degree of **sensitivity** of the original data used to create a data product rather than the data product itself, whose sensitivity is evaluated under the 'safe projects' principle. The level of sensitivity of an item of data must be evaluated and classified in accordance with Article 5 of the FADP.³²
- 3) **Safe people:** Is there a trust relationship with the recipient of the data product? If yes, of what kind and involving what level of trust? Trust should be interpreted here as the likelihood of the recipient of the data product attempting to extract identifying data. This trust relationship may take different forms, in particular professional secrecy, a contract, legal authority or accreditation.

³⁰ <https://www.bfs.admin.ch/bfs/fr/home/dscc/dscc.assetdetail.29325685.html>

³¹ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

³² https://www.fedlex.admin.ch/eli/cc/2022/491/en#art_5

For example, the various entities of the Federal Administration are bound by official secrecy, limiting the risk of identifying data being disclosed. In the case of collaborations outside the Federal Administration, a certain level of trust can be established through contractual measures. By contrast, any data product that takes the form of open government data places no restrictions on the recipient, who is therefore not bound by any trust relationship.

- 4) **Safe settings:** Does the technical, organisational and legal framework for data processing offer sufficient privacy guarantees? The 'safe people' principle allows a trust schema to be established formalising the use of PETs to offset the risks that are identified in points 1 and 3, and that create a potential lack of trust between the parties involved in creating the data product.
- 5) **Safe outputs:** Is the risk of identifying data being disclosed by the data product controlled? The use of output privacy enhancing technologies makes it possible to adapt the guaranteed limit on the risk of identifying data being disclosed, in terms of the level of sensitivity of the data, the level of trust in the recipient and the nature of the project.

The five principles can be grouped into two categories as shown in Figure 7.

- The principles corresponding to the risk evaluation for a project, responding to the question 'What?' and comprising principles 1 to 3, are comparable to the 'description of intended processing' and 'risk evaluation' components of a DPIA (Art. 22 FADP). The evaluation of this category is above all **subjective**, as it involves verifying the acceptability of the project in relation to social, cultural and political norms. If the level of risk for one of these principles is deemed to be too high, this may be prohibitive and justify the rejection of the project.
- The risk management measures, responding to the question 'How?', comprise the principles safe settings and safe outputs. In this category, which covers the final component of a DPIA, the risks of unauthorised access, usage limits and disclosure can be **objectively** evaluated and mitigated through organisational, legal and/or technical measures. It is in this category that PETs offer a **technical solution** for counteracting a potentially higher risk in the first category.

This type of **risk-based approach** is similar to existing regulations and requirements with regard to the application of existing legal bases, such as the data protection impact assessments required by the FADP³³ or the FSO application forms for data linkage.³⁴ This section thus illustrates how the PETs presented in this report may facilitate the implementation of this type of governance framework. They enable the risk of disclosure to be controlled for all types of identifying data (whether it relates to natural persons, legal entities, companies, households or groups of persons) and whatever the context for the secondary use of the data. They therefore appear to be a key element in the application of data science that not only preserves privacy but is also human-centric and trustworthy.

³³ <https://www.bj.admin.ch/bj/en/home/staat/datenschutz.html>

³⁴ <https://www.bfs.admin.ch/bfs/en/home/services/data-linkages.html>

7 Recommendation for the Federal Administration

To guarantee the availability of data, it is imperative that efforts to catalogue and above all harmonise data are continued. As well as promoting standardisation of the data collected, these initiatives also improve interoperability between different systems and databases, thereby facilitating access, secondary use by the various services and departments and the application of privacy control technologies.

It is also crucial to build up know-how in data science and privacy enhancing technologies within the Federal Administration. Professionals with expertise in this area are essential to analyse and interpret the vast quantities of data collected, while experts in privacy technology guarantee that sensitive data is protected against unauthorised access and privacy breaches. At the same time, an appropriate IT infrastructure must be implemented to permit prototyping and deployment of these technologies. This includes powerful servers, computing chips that permit the deployment of trusted execution environments, and the software tools necessary for testing and implementing the services required for the application of privacy preserving data science.

Confidential data science services such as those described in the two use cases set out in this report offer a direct response to one of the challenges currently facing the Federal Administration: in view of the growing demand for these services, which need to be provided on a large scale, notably in connection with the Digisanté project, the potential for the use of privacy enhancing technologies will enable the administration to simplify and accelerate their implementation and limit the resources required to provide them, while also guaranteeing the highest standards of privacy protection. These services are relevant not just for the FDHA but for the Federal Administration as a whole, which would justify them being made available to the entire administration.

A governance framework for the secondary use of data, such as the Five Safes model, also needs to be adopted to regulate the secondary use of the data held by the Federal Administration. This framework must seek to provide a structured approach based on evaluating the risks of a breach of fundamental rights – in particular the right to privacy as safeguarded by data confidentiality – and to describe the organisational, technical and legal measures implemented to limit these risks. It is therefore essential to establish a legal basis governing the secondary use of data regulated by this governance framework. Clear and precise legislation will allow the responsibilities and obligations of the different stakeholders to be defined, while at the same time providing a solid legal framework within which the Federal Administration can fulfil its mandate.

Finally, transparent communication regarding the technical and organisational measures implemented when using the data held by the Federal Administration in data science projects is vital to ensure the trust and cooperation of Swiss society as a whole. This transparency must set out in detail all measures permitting optimum implementation of the principle of privacy by design, in order to demonstrate the administration's commitment to preserving privacy.