BANK OF CANADA
BANQUE DU CANADA

Staff Discussion Paper/Document d'analyse du personnel—2025-1

Last updated: January 8, 2025

Privacy-Enhancing Technologies for CBDC Solutions

by Rakesh Arora, Han Du, Raza Ali Kazmi and Duc-Phong Le

Information Technology Services Department
Bank of Canada
RakeshArora@bankofcanada.ca, HDu@bankofcanada.ca

DOI: https://doi.org/10.34989/sdp-2025-1 | **ISSN** 1914-0568

Acknowledgements

The authors thank Lynne Graaskamp and Cyrus Minwalla for their insightful comments and constructive feedback on the first draft of the paper, which have significantly improved the quality of this work.

Abstract

With the rapid digitization of financial transactions, central banks have given considerable focus in recent years to the research and development of central bank digital currencies (CBDCs). While CBDCs could offer several advantages, there are concerns about end-user privacy. Traditional methods of protecting confidentiality in banking and financial systems have primarily relied on data encryption and access control techniques. However, these techniques alone are inadequate, especially in cases where data are shared across different entities because privacy in such situations is typically governed by legal frameworks. Privacy-enhancing technologies (PETs) can offer robust protection for data throughout their lifecycle, whether stored, in transit or during processing, and ensure privacy is maintained even when data are extensively shared or analyzed. This study explores the use of PETs in the design of CBDC systems, potentially paving the way for solutions that better safeguard end-user privacy and meet rigorous data protection standards. While PETs promise significant advancements in privacy protection, they present some challenges in implementation. They can introduce performance overheads and add complexity to systems, and their effectiveness and applicability are currently limited due to their early stage of development. As these technologies evolve, it is crucial for organizations to carefully consider these factors to fully leverage PET benefits while managing associated challenges. This paper provides a comprehensive overview of how PETs can transform privacy design in financial systems and the implications of their broader adoption.

Topics: Central bank research, Digital currencies and fintech, Financial system regulations and policies, Payment clearing and settlement systems

JEL codes: E, E4, E42, O, O3, O31

Résumé

La numérisation rapide des transactions financières a amené les banques centrales à s'intéresser de près aux études consacrées aux monnaies numériques de banque centrale (MNBC) et à leur conception. Ces monnaies auraient plusieurs avantages, mais elles soulèvent des questions par rapport à la protection des données personnelles des usagers. Les méthodes généralement employées pour protéger la confidentialité des données des systèmes bancaires et financiers reposent tout d'abord sur e chiffrement de données et des dispositifs de contrôle de l'accès. Seuls, ces dispositifs sont toutefois inadéquats, tout particulièrement dans les situations où plusieurs entités s'échangent des données car, dans ces cas, la protection des données personnelles est encadrée normalement par des cadres juridiques. Les technologies d'amélioration de la confidentialité peuvent sécuriser les données tout le long de leur cycle de vie, que ces données soient stockées, en transit ou en traitement. Ces techniques permettent de préserver la confidentialité même lorsque les données sont fréquemment échangées ou analysées. Notre étude porte sur l'usage de telles techniques dans la conception des systèmes reposant sur les MNBC. Elle pourrait ouvrir la voie à des solutions qui améliorent la protection

de la vie privée et répondent à des normes strictes dans le domaine de la protection des données. Bien que des progrès majeurs puissent découler des technologies d'amélioration de la confidentialité, le déploiement de telles techniques demeure difficile. Elles peuvent allonger le temps-système et complexifier les systèmes. Par ailleurs, leur efficacité et leur applicabilité sont pour le moment limitées, car ces techniques sont au tout premier stade de développement. À mesure qu'elles évolueront, il sera déterminant de prendre en compte les complications qui leur sont associées si les organisations veulent tirer pleinement profit de leurs retombées. Notre étude dresse un tableau complet de la manière dont ces techniques peuvent transformer la conception de la protection des données des systèmes financiers et présente les conséquences de leur adoption accrue.

Sujets : Recherches menées par les banques centrales, Monnaies numériques et technologies financières, Réglementation et politiques relatives au système financier, Systèmes de compensation et de règlement des paiements

Codes JEL : E, E4, E42, O, O3, O31

1 Introduction

In an era marked by the rapid digitization of financial transactions and the declining use of cash, central bank digital currencies (CBDCs) have emerged as a focal point of research and development. This shift toward digital payments is accompanied by the proliferation of blockchain, cryptocurrencies and stablecoins, posing both opportunities for and threats to the financial landscape. As central banks, including entities like the Bank for International Settlements, engage in extensive research and development of CBDCs, it becomes imperative to address the escalating privacy concerns associated with these advancements.

Privacy, broadly defined, encompasses the right of individuals and entities to control their personal information, ensuring it is collected, used and shared in ways that respect their autonomy and safeguard against unwanted disclosure or exploitation. Privacy concerns in the context of CBDCs are multifaceted, involving apprehensions of end users and merchants alike. The interconnected nature of our digital world has led to an increased collection of personal information, making it crucial for individuals and businesses to manage privacy risks and protect against unauthorized access and data misuse. Central banks, entrusted with the responsibility of introducing CBDCs, face challenges in balancing the imperative for privacy with the compliance demands of a digital financial landscape.

Existing privacy laws and regulations, such as the General Data Protection Regulation (European Parliament and Council of the European Union 2016) and the *Personal Information Protection and Electronic Documents Act* (2000), provide a fundamental framework for safeguarding individuals' data. However, as the digital environment evolves, accompanied by a proliferation of sensitive data, these laws are evolving to address emerging challenges. Designers of CBDCs must take a proactive approach that prioritizes user-data protection rights from the outset. This involves handling sensitive user data and necessitates a privacy-by-design approach, including integrating privacy-enhancing technologies (PETs) and embedding privacy considerations within the architecture to ensure the protection and confidentiality of user information. Such an approach not only ensures compliance with existing regulations but also anticipates and addresses emerging concerns.

Lately, PETs have emerged as a crucial means to address privacy concerns associated with CBDCs. A CBDC designed with PETs could minimize personal data exposure and maximize data integrity and confidentiality. While no consensus exists on a common definition of PETs, in this paper we investigate a diverse set of technologies that preserve the confidentiality of transactions and mitigate the risks posed by increased data collection and cyber threats. We further classify and describe candidate privacy solutions for a CBDC in Section 2. As central banks explore the integration of these PETs into the design of CBDC systems, understanding the implications and benefits of such technological advancements becomes essential.

Potential PETs that are applicable in digital payments are broad, covering cryptographic, statistical and procedural technologies. Asrow and Samonas (2021) and the Bank of England (2023) have summarized existing PETs that could potentially be used in the design of a CBDC system. The blockchain industry has seen the implementation of many cryptographic PET technologies to protect the confidentiality of senders, receivers and transaction amounts. For instance, Monero (van (Saberhagen 2013) has implemented a ring confidential transactions protocol consisting of ring signatures (Rivest, Shamir, and Tauman 2001) and a Pedersen commitment (Pedersen 1992); Zcash (Hopwood et al. 2021) has implemented zero-knowledge proofs (ZKPs) (Ben-Sasson et al. 2018) to provide transaction confidentiality; and the Swiss National Bank and the Bank for International Settlements (2023) have explored the feasibility of blind signatures (Chaum 1983) in CBDC design.

The primary inquiry of this research focuses on using PETs within the framework of CBDC design to protect consumers' personal data while simultaneously addressing the imperative of regulatory compliance. Consequently, this paper introduces a CBDC design paradigm to explore the use of cutting-edge PETs in providing high levels of privacy. The objective of the system is to give consumers control over their personal data in the CBDC system, striking a delicate balance between user privacy expectations and the demands of regulatory frameworks concerning anti-money laundering (AML) and counter-terrorism financing. We first present a comprehensive and systematic description of PETs that could be applied to digital currencies. Then we unveil a privacy-centric CBDC design framework encompassing key components such as user onboarding, identity and access management, transaction processing, regulatory compliance, data analytics and digital wallets. We conduct an in-depth analysis of privacy objectives for each component, followed by an investigation into the possible integration of PETs within the design of each component. Furthermore, we identify and address the inherent challenges associated with incorporating PETs into the proposed CBDC design.

The rest of the paper is organized as follows. Section 2 briefly summarizes existing PET technologies

that could be applied in digital currencies. Section 3 discusses a potential CBDC design with privacy. The subsequent sections consider specific components in a CBDC system, including user onboarding in Section 4, identity and access control in Section 5, value transportation and transaction processing in Section 6, compliance in Section 7, data analytics in Section 8 and wallets in Section 9. For each component, we start by briefly describing its functionalities, followed by a discussion of the privacy objectives required for that specific component. Next, we introduce potential PETs that could be used and discuss the challenges associated with incorporating them into the design and implementation. Finally, we present conclusions and open research questions in Section 10.

2 Taxonomy of privacy-enhancing technologies

PETs play an essential role in balancing the benefits of data-driven technologies with the protection of individual data. Many PET techniques have been introduced over time, along with the development of digital systems and the expansion of data privacy regulations. The first principle for preserving individuals' privacy is minimizing the information collected, processed and stored. Data anonymization (Dalenius 1986; Samarati and Sweeney 1998a) and pseudonymization (Samarati and Sweeney 1998b) are two of the first approaches to provide privacy for published data. These techniques alter data and aim to break the link between original data and published data. Encryption for data in rest and transit is quite prevalent in modern digital systems. Ideas for protecting data during processing using cryptographic solutions were introduced in the 1970s and 1980s. These ideas included computing on encrypted data (Rivest, Adleman, and Dertouzos 1978), secure multiparty computation (SMPC) (Yao 1982), ZKPs (Goldwasser, Micali, and Rackoff 1989) and many other cryptographic primitives.

We cover a broad range of software and hardware technologies in this paper. Most of these technologies use advanced cryptography to provide data protection. While the techniques are not fundamentally new, the exploration of new applications is gaining momentum in academia and industry. The following are some of the main contributing reasons for the recent uptake in interest in PETs:

- **Distributed ledger technology (DLT)/blockchain systems**: These systems provide weak privacy since the ledger is available and visible to many entities. Many projects (e.g., Zcash, Aztech) and research are ongoing to apply PETs for better data protection in systems based on distributed ledger technology.
- **Digital money and digital payments**: In response to the decline in cash use and the emergence of cryptocurrencies, central banks are working on the potential digital form of countries' fiat currency or CBDC (Mikhalev et al. 2021). Similar to cash, CBDCs should provide a high level of privacy to consumers in their technical design. In the payment industry, while customer data must be privately protected, it is important to follow AML regulations as well as those combating the financing of terrorism to prevent illicit usage. A PET must satisfy both privacy and these latter regulations.
- **Monetizing of data**: Data owners are gathering vast amounts of data and seeking innovative business models to monetize this data while adhering to laws and regulations. These models often involve sharing data with other entities, which introduces many privacy challenges. PETs can be crucial in addressing these challenges by facilitating secure data sharing that protects user privacy.
- **Strict privacy regulations**: Awareness about privacy and strict privacy regulations (e.g., the General Data Protection Regulation) have motivated many to find innovative solutions to design privacy-centric systems.

2.1 Classification of existing privacy-enhancing technologies

PETs can be classified according to different criteria, such as privacy controls (e.g., governance controls, data minimization or statistical disclosure controls); cryptographic solutions versus non-cryptographic solutions; or their technical designs. The literature has introduced several PET classifications (Seničar, Jerman-Blažič, and Klobučar 2003; Montjoye et al. 2015; Privacy Commissioner of Canada 2017; Asrow and Samonas 2021). In this section, we present a taxonomy for PETs based on the classification system introduced by the US Federal Reserve Bank of San Francisco (Asrow and Samonas 2021).

- **Altering data**: These techniques primarily alter data to enhance privacy by breaking the association between individuals and their data. Some of these techniques include anonymization, data masking, differential privacy and synthetic data. This class of PETs can be used in statistical disclosure of data, such as for data analytics.
- **Shielding data**: These techniques focus on hiding the data from different entities or systems. They include encryption techniques, special signatures, ZKP and privacy-enhancing hardware. This class of measures could be used to protect the privacy of data when they are at rest, in use or in transit.
- **System and architecture**: This class of techniques focuses on organizing and processing data. Techniques include SMPC, data dispersion, privacy-preserving identity, access control and federated learning.

Table 1 lists the existing PETs based on the above classification. We briefly outline further technical details of these PETs in the remainder of this section.

Table 1: Classification of existing privacy-enhancing technologies

PETs		
Altering data	**Shielding data**	**System and architecture**
Data suppression	Special signatures	Multiparty computation
Anonymization	Homomorphic encryption	Data dispersion
Pseudonymization	Cryptographic commitments	Privacy-preserving digital identity
Synthetic data	Zero-knowledge proof	Federated learning
Differential privacy	Privacy-enhanced hardware	Onion routing
	Transaction tumbler	

2.2 Altering data

2.2.1 Data suppression

This technique provides data protection by deleting or removing information (columns, rows or specific records) that is not required in the shared data. Two types of data suppression exist:

- **Attribute suppression** refers to the removal of a column of data in a dataset when this column is not required in the anonymized dataset or when it cannot be suitably anonymized with another technique. This technique should be applied at the start of the anonymization process, as it is an easy way to decrease identifiability at this point.
- **Record suppression** refers to the removal of a row in the dataset. This technique affects multiple attributes at the same time. It aims to eliminate *outlier records* that are unique or easily re-identifiable or that do not fulfill criteria such as k-anonymity. This technique can be applied before or after other techniques (e.g., generalization) have been applied.

2.2.2 Anonymization

Data anonymization (Dalenius 1986), a classical PET used in publishing data, is the process of protecting sensitive information by removing direct or indirect identifiers from a dataset. Anonymizing data can help avoid compliance risks with privacy regulations (such as the *Personal Information Protection and Electronic Documents Act* and General Data Protection Regulation) by ensuring that personal information is not traceable to individuals. According to paragraph 26 of the preamble to the General Data Protection Regulation (European Parliament and Council of the European Union 2016): "The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable." Various techniques are available for anonymizing the user data:

- **Data perturbation**: This method protects data by incorporating "noise" to make personally identifiable information (PII) unlinkable for unauthorized users. There are generally two noise-adding techniques: *additive noise* (R. Agrawal and Srikant 2000; D. Agrawal and Aggarwal 2001) and *multiplicative noise* (Liu, Kargupta, and Ryan 2006). Although additive noise techniques can be mitigated by certain signal-processing methods (Kargupta et al. 2003), multiplicative perturbation techniques such as those based on random projection can circumvent this problem (Liu, Kargupta, and Ryan 2006).
- **Data permutation**: This technique splits the data records into several groups. It then shuffles the values of the sensitive attributes within each group, de-associating them from the identifiers within each group. This technique can provide high privacy, but may produce inaccurate analysis.

The data anonymization method has been shown to be relatively insecure when multiple data sources are merged. With the rapid expansion of digital information, along with advancements in machine learning models and data mining tools, attackers can more easily extract personal information. Even when identifiers are removed from data, attackers can employ de-anonymization techniques to reverse the data anonymization process, as data are often shared across multiple sources. This allows attackers to cross-reference these sources and uncover personal information. (See paper by Al-Azizy et al. (2016) for more details of de-anonymization approaches.)

2.2.3 Pseudonymization

Pseudonymization involves replacing identifiable data with fabricated values, a technique also known as *coding*. Pseudonyms can be either *irreversible* or *reversible* (by the owner of the original data). In the case of irreversible pseudonymization, the original values are permanently discarded and the process is non-reversible. Conversely, with reversible pseudonymization, the original data are securely stored and can be re-associated with the pseudonym if necessary. This capability sets pseudonymization apart from anonymization. Various methods can be used in pseudonymization:

Generalization Generalization entails substituting individual attribute values with broader categories (Samarati and Sweeney 1998b). For instance, this can involve converting a specific age into an age range, or a precise location into a more general one. The purpose of generalization is to obscure demographic information to meet privacy standards while still enabling meaningful data analysis. Generalization strategies can be classified into two categories: *global* and *local*.

- In global generalization, a given value for a given column will always be generalized in the same way. For example, if you decide to transform age 34 into age range 30–35 for one record, all records that have ages between 30 and 35 will be transformed into this fixed range.
- Local generalization does not have this constraint. It allows you to pick a different generalization for each record. For example, a value 34 in the age column might stay untouched for one record and be generalized for another.

Data masking Data masking or data obfuscation, involves altering the original data with modified content, such as different characters or other data types. This technique can protect sensitive information by making it unreadable and useless to unauthorized viewers, without affecting the data's usability for legitimate processing and analysis. (Hush Hush, nd). Data involved in any data-masking or obfuscation must remain meaningful on several levels:

- The data must still be meaningful for the application's logic. For instance, if elements of addresses, such as cities and suburbs, are replaced with alternative names, any features within the application that validate postal codes must continue to function correctly and as intended.
- The data should be sufficiently altered to ensure that it is not apparent that the masked data originate from a production data source.
- When multiple databases within an organization contain the specific data element being masked, the masked values may need to be consistent across all these databases.

The techniques available for data masking include substitution, shuffling and applying variances to numbers and dates.

Tokenization Tokenization is a process that replaces a sensitive data element with a token. In payment systems, tokens—usually random strings of numbers—replace the 16-digit primary account number of payment cards. These tokens can be device-based, as seen in Apple Pay, or cloud-based, as used in Google Pay. Entities called token server providers will perform tokenization in a secure environment and hold a key that allows tokens to be securely matched to the true primary account number they represent. This approach has become standard practice in the card payment industry (Stapleton and Poore 2011) to protect both consumers and businesses from the risks associated with unauthorized access to sensitive payment information. Single-use tokens are generated for specific transactions, adding an extra layer of security to payment transactions.

2.2.4 Synthetic data

Unlike anonymization methods, these techniques are employed primarily to generate new datasets rather than altering the existing dataset. They do this through learning from the original dataset and using this information to create new independent samples. Synthetic data ideally retain the statistical patterns, properties, features and characteristics of the original data not limited to format and relationships among attributes. This technique is best employed when a large amount of data is required for system testing but use of the actual data is limited or prohibited. The synthetic data are broadly classified into three categories:

- **Fully synthetic data**: These data are completely synthetic and do not contain original data. To achieve this, the fully synthetic data generators first identify the density function of attributes in the original data and estimate the parameters of these density functions. Then, for each attribute, privacy-protected series are generated by randomly picking up the values from the estimated density functions.
- **Partially synthetic data**: In contrast to the fully synthetic data, the method used to generate partially synthetic data replaces only values of the selected sensitive attribute with synthetic values. The original values are replaced only if they possess high risk of confidentiality. Disclosure risk is higher in partially synthetic data than in fully synthetic data because the former contain original data along with imputed synthetic data.
- **Hybrid synthetic data**: These data are generated using both original and synthetic data. For each record of original data, a nearest record in the synthetic data is chosen, and both are combined to form hybrid data. The hybrid synthetic data hold the advantages of both fully and partially synthetic data.

2.2.5 Differential privacy

The concept of differential privacy was introduced by Dwork et al. (2006). An algorithm is considered differentially private if an observer, upon seeing its output, cannot determine whether a specific individual's information was included in the computation. Differential privacy arises frequently in discussions about pinpointing individuals whose data might be present in a database. While this concept does not explicitly address identification and reidentification attacks, algorithms that are differentially private are likely to withstand such attacks. Differential privacy offers strong and robust guarantees that facilitate modular design and analysis of differentially private mechanisms due to its composability, robustness to post-processing, and graceful degradation in the presence of correlated data. Dwork et al.'s paper presents both a mathematical definition of differential privacy and a mechanism based on the addition of Laplace noise that satisfies the definition as follows:

Definition of ϵ-differential privacy: Let ϵ be a positive real number and $\mathcal{A}$ be a randomized algorithm that takes a dataset as input (representing the actions of the trusted party holding the data). Let $\mathrm{im}(\mathcal{A})$ denote the image of $\mathcal{A}$. The algorithm $\mathcal{A}$ is said to provide ϵ-differential privacy if, for all datasets D_1 and D_2 that differ on a single element (i.e., the data of one person), and all subsets S of $\mathrm{im}(\mathcal{A})$:

$$\Pr[\mathcal{A}(D_1) \in S] \leq e^{\epsilon} \times \Pr[\mathcal{A}(D_2) \in S],$$

The Laplace mechanism: The Laplace mechanism adds Laplace noise (i.e., noise from the Laplace distribution), which can be expressed by probability density function $\mathrm{noise}(y) \propto \exp(-|y|/\lambda)$, having mean zero and standard deviation $\sqrt{2}\lambda$. Now, in our case we define the output function of $\mathcal{A}$ as a real valued function (called the transcript output by $\mathcal{A}$) as $\mathcal{T}_{\mathcal{A}}(x) = f(x) + Y$, where $Y \sim \mathrm{Lap}(\lambda)$ and f is the original

real valued query/function we plan to execute on the database. Now clearly $\mathcal{T}_{\mathcal{A}}(x)$ can be considered to be a continuous random variable, where

$$\frac{\text{pdf}(\mathcal{T}_{\mathcal{A},D_1}(x) = t)}{\text{pdf}(\mathcal{T}_{\mathcal{A},D_2}(x) = t)} = \frac{\text{noise}(t - f(D_1))}{\text{noise}(t - f(D_2))},$$

which is at most $e^{\frac{|f(D_1)-f(D_2)|}{\lambda}} \leq e^{\frac{\Delta(f)}{\lambda}}$. We can consider $\frac{\Delta(f)}{\lambda}$ to be the privacy factor ϵ. Thus, $\mathcal{T}$ follows a differentially private mechanism (as can be seen from the definition above). Though we have used Laplacian noise here, other forms of noise, such as the Gaussian noise, can be used, but they may require a slight relaxation of the definition of differential privacy.

In payment systems, differential privacy can be applied to enhance the confidentiality of transaction data. For example, a financial institution might use a differential privacy algorithm to add noise to transaction datasets, such as the transaction amounts or timestamps. This modification makes it impossible to identify individual transactions or link them back to specific customers while still allowing the institution to analyze overall spending patterns and trends.

2.3 Shielding data

2.3.1 Special digital signatures

While the aim of using digital signatures is to maintain authenticity of messages and non-repudiation, special signatures exist to enhance privacy by, for example, hiding the signer in a group of users (e.g., group signatures, ring signatures) or hiding the content of a signed message (e.g., blind signatures). This section will briefly recall some special signature schemes.

Blind signatures The first special digital signature scheme introduced to the payment system as a means of privacy is the blind signature. It was introduced for untraceable payments by Chaum (1983). The blind signature is a form of digital signature in which the message is disguised (blinded) before it is signed, that is $m' \equiv mr^e \bmod N$. The blind signature $s' \equiv (m')^d \bmod N$, when given to the intended recipient, will have the necessary information to remove the blinding factor to reveal the resulting signature $s = s' \cdot r^{-1} \bmod N$. The signature s then can be publicly verified against the original, unblinded message in the manner of a regular digital signature. The potential applications of Rivest, Shamir and Adleman's (RSA) blind signatures led to a standard recently published by the Internet Engineering Task Force (Denis, Jacobs, and Wood 2023).

The initial idea of blind signatures is for untraceable payments (Chaum 1983). Further research has expanded its applications to include anonymous authentication. A user possessing valid blind signatures can access a server while maintaining anonymity, meaning the service provider cannot identify the actual user accessing the service. Additionally, it ensures unlinkability, preventing the provider from associating a previously issued token with a specific user.

Group signature scheme A group signature scheme, introduced by Chaum and van Heyst (1991), is a type of signature scheme in which an entity called a group manager owns a master key. When a new user joins the group and is onboarded, they receive a signing key. A user then can sign messages anonymously in an unlinkable fashion. The group manager is the only entity that can learn the identity of the user from the signature. Moreover, the group manager and other members cannot forge a signature on behalf of a non-participant member.

A group signature offers nice privacy properties, including the anonymity of signers and unlinkability—such that given two messages and their signatures, we are unable to determine whether the signatures originate from the same signer. The group signature scheme and its variants have a number of privacy-preserving applications, such as e-commerce systems, auctions, trust computing groups and vehicle safety communication (Garms 2020). A group signature scheme can be either static (Chaum and van Heyst 1991), where membership in the group is decided at the time a group is set up and new members cannot be added or removed, or dynamic, where group members can be added or removed (Delerablée and Pointcheval 2006; Ling et al. 2017).

Ring signatures Ring signatures were first introduced by Rivest, Shamir, and Tauman (2001). These allow an individual to anonymously sign a message on behalf of a group, without revealing their identity. Unlike group signatures, ring signatures do not require a setup phase. The signer does not need the knowledge, consent or assistance of other members of the ring to include them. Instead, all that is required

is knowledge of their regular public keys. Ring signatures offer unconditional anonymity—that is, untraceability and unlinkability. Untraceability makes it impossible to identify the signer, while unlinkability makes it impossible to determine if two signatures were produced by the same user. This level of anonymity ensures robust privacy protection for individual signing actions. However, this feature could also be exploited for illegal activities. While all group, blind and ring signatures offer signer anonymity, only ring signatures and their variants are used in blockchain networks because they do not require a trusted setup phase. This makes ring signatures more adaptable and secure for use in decentralized environments like blockchains. Noether (2015)presents a combination of three techniques: (1) ring signatures to protect the sender's privacy, (2) homomorphic commitment to hide the transaction amount and (3) stealth addresses to protect the recipient's privacy to a protocol. Ring confidential transactions, or RingCT, is the protocol responsible for implementing private transactions in Monero.

One-time (ring) signatures A one-time signature scheme is like a signature scheme except that it is secure as long as a secret key is used to sign only a single message. The signing key can be recovered if it is used more than one time. The first one-time signature scheme was introduced by Lamport (1979). Its security is based on one-way functions and allowed the signing of one message per pair of private/public keys. van Saberhagen (2013) combines the ideas of one-time signatures and ring signatures to propose a new primitive: that is, one-time ring signatures. Like one-time signatures, the signer in this signature scheme uses the private key to sign only one time on behalf of a group.
In contrast to one-time signatures, which do not inherently prevent double spending on blockchains, the `Link` algorithm in one-time ring signatures enables the linking of two valid signatures, σ_1 and σ_2, to determine whether they were signed by the same private key. More details on this can be found in the CryptoNote whitepaper (2013).

2.3.2 Homomorphic encryption

Homomorphic encryption (HE) is a form of encryption that allows computation on ciphertexts, generating an encrypted result that, when decrypted, matches the result of the operations as if they had been performed on the plaintext. The computations are represented as either Boolean or arithmetic circuits. This enables sensitive data to be processed securely without exposing it, thereby preserving privacy throughout the computation process. The initial idea of computing over encrypted data was first introduced by Rivest, Adleman, and Dertouzos (1978). More formally, M and C are *plaintext* and *ciphertext* spaces, respectively. An encryption scheme HE is called homomorphic over an operation ''$*$" if it supports the equation $HE(m_1) * HE(m_2) = HE(m_1 * m_2)$, for $\forall\, m_1, m_2 \in M$.

The operation $*$ could be an addition or multiplication. Given $c_1 = HE(m_1)$ and $c_2 = HE(m_2)$, the homomorphic properties of a fully homomorphic encryption scheme can be shown as follows:

Homomorphism over addition:

$$c_1 + c_2 = HE(m_1) + HE(m_2)$$

Homomorphism over multiplication:

$$c_1 \cdot c_2 == HE(m_1) \cdot HE(m_2).$$

To evaluate an arbitrary function, it is sufficient that a homomorphic encryption scheme supports both addition and multiplication operations, as addition and multiplication are functionally complete sets over finite sets. HE schemes that support only one of these two operations—either addition or multiplication—on encrypted data are called partially homomorphic encryption (PHE). RSA (Rivest, Shamir, and Adleman 1978) is a typical examples of PHE. Otherwise, schemes supporting both operations are called fully homomorphic encryption (FHE). The first FHE scheme was not realized until 2009, when Gentry (2009) introduced a construction-based ideal lattice scheme. Although this scheme was very inefficient due to its costly bootstrapping operation, it inspired a number of subsequent works (Brakerski, Gentry, and Vaikuntanathan 2012; Chillotti et al. 2017). The new form of FHE, Torus-FHE (Chillotti et al. 2020), introduced programmable bootstrapping, allowing for the computation of more than thousands of bootstrapping operations per second. This scheme led to huge performance advantages, making FHE practical today.

2.3.3 Cryptographic commitment

A cryptographic commitment scheme allows us to commit to a chosen message while preserving its secrecy (with the ability to reveal it later) by publishing its hash value. A binding factor can be used when data size is small to prevent a brute-force attack. A commitment $Com(m, r)$ to message m and a blinding factor r has the following properties:

- **Hiding**: One party wants to commit the message m without revealing the content of m itself.
- **Binding**: If one party makes a commitment to m, they reveal a different message m' instead.

Brassard, Chaum, and Crépeau (1988) provide a formal definition of a commitment scheme, as follows:
A commitment scheme consists of three polynomial-time algorithms $C = (\texttt{Gen}, \texttt{Commit}, \texttt{Open})$ satisfying the following constraints.

- $r \leftarrow \texttt{Gen}(1^\lambda)$: take a security parameter 1^λ and output a random number r.
- $c \leftarrow \texttt{Commit}(m, r)$: given a message m and randomness r, compute as output a value c that hides message m and such that it is computationally impossible to compute any pair (m', r') such that $\texttt{Commit}(m', r') = \texttt{Commit}(m, r)$.
- $b \leftarrow \texttt{Open}(c, m, r)$: given (c, m, r) with a commitment c, a message m and randomness r, the algorithm returns true if and only if $c = \texttt{commit}(m, r)$.

Pedersen commitment The Pedersen commitment is a commitment scheme that is binding under a discrete logarithm assumption (Pedersen 1992). Given an elliptic curve E defined over a finite field $GF(p)$, assume that E has a group of points of large order q in which the discrete logarithm is hard, and two random public generators g and h. The commitment of a message m is a point c on the elliptic curve E that binds a message m and a random r to a point c on E. The Pedersen commitment is defined as follows:

$$Com(m, r) = g^m h^r$$

It would be infeasible to calculate another pair m', r' that can produce the same commitment $Com(m)$. The Petersen commitment is used in cryptocurrencies such as Monero to keep the amount of transactions confidential. Additionally, Pedersen commitments are additively homomorphic, meaning the sum of a set of commitments equals a commitment to the sum of the data, with the binding factor set as the sum of the individual binding factors.

2.3.4 Zero-knowledge proofs

Goldwasser, Micali, and Rackoff (1989) first put forward and analyzed the concept of interactive proof systems. This led to the creation of an important branch of cryptography and computational complexity theory: that is, ZKP. In a ZKP system, a *prover* convinces a *verifier* that some statement is true or some computations were correctly executed while leaking nothing but the validity of the assertion. ZKP has a broad application because of this nature. As a simple example, consider the case where the prover claims to have a way of factorizing large numbers. The verifier will send the prover a large number and the prover will send back the factors. Successful factorization of several large integers will decrease the verifier's doubt about the truth of the prover's claim. In spite of this, the verifier will learn nothing about the factorization method.

Formally, a ZKP model is defined as an interactive proof system (P, V), where P is a prover and V is a verifier. Protocol (P, V) is for proving a language membership statement for a language over $\{0, 1\}$. Let L be a language over $\{0, 1\}^*$, for a membership instance $x \in L$, P and V must share the common input x. A proof instance is denoted as $(P, V)(x)$. Upon completing the interaction between two parties, the output of the protocol should be of form $(P, V)(x) \in \{\textbf{Accept}, \textbf{Reject}\}$, representing V's acceptance or rejection of P's claim that $x \in L$. Note that an interactive ZKP system can be converted to a non-interactive ZKP system by using the Fiat-Shamir transformation (Fiat and Shamir 1987).

A ZKP must have the following properties:

- **Completeness**: If the statement is correct, then the verifier will "always" accept.
- **Soundness**: If the statement is incorrect, then the verifier will "always" reject.

- **Zero knowledge**: No (malicious) verifier can get any extra information beyond $x \in L$ from the proof procedure, except the correctness of the statement (Goldwasser, Micali, and Rackoff 1989).
- **Succinctness**: Both the length of the proof and the verification time are bounded by polylog functions with respect to the size of the circuit C representing $x \in L$ (Kilian 1992).
- **Proof of knowledge (POK) or argument of knowledge (AOK)**: Extracting a witness from an acceptable proof/argument would be efficient (Blum, Feldman, and Micali 1988).

A succinct non-interactive argument that possesses the properties of zero-knowledge and proof of knowledge is referred to as *zk-SNARK*. zk-SNARK protocols can require either an initial trusted set-up (per circuit (Groth 2016) or universal (Gabizon, Williamson, and Ciobotaru 2019)) or in zk-STARKs a transparent set-up (no trusted set-up (Ben-Sasson et al. 2018)). A universal trusted set-up would be better than a trusted set-up per circuit, and a transparent set-up is even better than the universal set-up. This is because it does not use secret data, and thus anyone could verify that it ran correctly. Furthermore, zk-STARKs rely on collision-resistant hash functions, and thus it eliminates the number-theoretic assumptions of zk-SNARKs, which are computationally expensive and theoretically can be prone to attack by quantum computers.

Recently, ZKP has been attracting a great deal of attention because of its applications in blockchains. (Hopwood et al. 2021) implement zk-SNARKs, allowing a transaction to be verified without revealing any details about the transaction itself. Data are obscured using zk-SNARKs and recorded on the Zcash network, similar to other cryptocurrencies. Zcash's proofs are both succinct and non-interactive, meaning that the proofs could be verified without requiring communication between the sender and verifier. Other applications include zk-rollup (Polygon Labs 2024), [1] which enhances the scalability of the blockchains, and zkBridge (Xie et al. 2022), which provides security and privacy for cross-chain transactions over two different blockchain platforms.

2.3.5 Privacy-enhanced hardware

The simplest example of privacy-enhanced hardware is the use of trusted execution environments (TEEs) to perform secure computation on sensitive data. These technologies include Intel's Software Guard Extension (Intel® SGX) (Intel Corporation n.d.) and ARM's TrustZone (ARM n.d.). In a TEE, a software program is executed securely in an isolated and secure area. Data and programs are encrypted outside of the microprocessor and can be decrypted and executed with the cryptographic key only inside the microprocessor. Different from other cryptographic solutions, this technology allows any arbitrary programs to be executed securely.

A potential security issue associated with TEE technology is side-channel analysis (Kocher 1996), which can lead to information leakage. Given the variety of the potential privacy-enhanced hardware available for implementation, we also need to consider the vulnerabilities of each solution independently. Thus, to protect data confidentiality, a TEE must be securely designed and implemented. Today, a combination of hardware-based TEEs, software, and the infrastructure to support them is offered as a complete product, commonly referred to as confidential computing (Confidential Computing Consortium 2022). Numerous vendors have formed a consortium (Linux Foundation 2023) to collaborate on this initiative, emphasizing data protection during processing. This has become increasingly pertinent with the widespread shift from on-premises systems to cloud computing.

2.3.6 Transaction tumbler

Over the development of cryptocurrencies, several designs and improvements have been made to transaction tumblers, popularly known as "mixers." Originally, tumblers involved a scheme where multiple users would combine their unspent transaction output (UTXO) funds and spend to destination addresses in a bundle, thus providing some plausible deniability in terms of who intended to send what funds to which recipients (Bitcoin Wiki n.d.). Later on, better schemes were invented that allowed users to tumble their funds in zero-knowledge while hiding the transaction graph available from the original scheme. The most infamous implementation of this scheme is Tornado Cash.

Tornado Cash operates by enabling users to deposit funds of a fixed denomination into the protocol. Using a fixed denomination is crucial because varying amounts can increase the likelihood of correlating

[1] Polygon zkEVM offers an example of a zk-rollup application.

senders and receivers within this specific scheme. When users make a deposit in the system, a secret is simultaneously generated for a UTXO, allowing the original sender to transfer funds to a recipient without revealing their identity. This is because the scheme only verifies the presence of a valid secret, without disclosing which UTXO it is linked to. When the sender decides to move the funds, they can either (1) construct a ZKP that includes proof of transaction validity, recipient details, a nullifier hash and transaction fees and send it to a relayer who will then submit the transaction details to the on-chain smart contract on the sender's behalf, or (2) pass the secret directly to the recipient through an external channel, allowing the recipient to withdraw the funds themselves.

In this scheme, relayers are essential when the recipient lacks the funds to cover fees: for example, gas in Ethereum. The sender cannot execute the withdrawal transaction themselves, as it would compromise the privacy protections (Pertsev, Semenov, and Storm 2019). The latest version of the protocol now includes shielded transfers, which enable users to securely move pledged funds to other users within Tornado Cash and also support withdrawals of arbitrary amounts.

2.4 System and architecture

2.4.1 Secure multiparty computation

SMPC (also known as secure computation, multiparty computation or privacy-preserving computation) aims to create methods for parties to jointly compute a function over their inputs while keeping those inputs private. For example, Alice, Bob and Charlie, with respective inputs x, y and z denoting their salaries, want to find the highest of the three salaries without revealing to each other how much each of them makes. Mathematically, this translates to them computing:

$$F(x, y, z) = max(x, y, z)$$

Unlike traditional cryptographic tasks, where cryptography ensures security and integrity of communication or storage and the adversary is outside the system of participants (an eavesdropper on the sender and receiver), the cryptography in this model protects participants' privacy from each other. The foundation for SMPC was introduced in (Yao 1982) and (Goldreich, Micali, and Wigderson 1987) with work on mental poker. This cryptographic work simulates game playing/computational tasks over distances without requiring a trusted third party. In this setting, the computation is carried out interactively between several participating parties in such a way that sensitive data are kept hidden (e.g., encrypted or shared among protocol participants) and only the desired output of the computation is available.

An SMPC protocol must offer the following security properties:

- **Input privacy**. No information about the private data held by the parties can be inferred from the messages sent during the execution of the protocol.
- **Independence of inputs**: Corrupted parties must choose their inputs independently of the honest parties' inputs. This property is crucial in a sealed auction, where bids are kept secret and parties must fix their bids independently of others.
- **Correctness**: Any proper subset of adversarial colluding parties willing to share information or deviate from the instructions during the protocol execution should not be able to force honest parties to output an incorrect result.

In the following sections, we show that many operations in a CBDC system could be deployed over an SMPC system in which participants are end users, central banks and commercial banks.

Another interesting use case of SMPC is to prevent a single point of failure. A party splits its data into many different shares stored in multiple servers, enhancing not only the resilience but also the security of data. A typical example is SMPC-based digital wallets (Zengo Ltd. n.d.), in which the private signing key is sharded into n pieces and distributed to n parties using a secret sharing scheme (e.g., Shamir (1979) secret sharing). It then requires k parties (Shamir's scheme requires $k \geq 2n/3$) involved in the process of recovering the private key.

2.4.2 Data dispersion

Privacy can be enhanced by incorporating data-dispersion techniques, which apply encryption methods to all targeted data. A widely used method is cryptographic splitting, in which data is fragmented, encrypted

and distributed in a manner that prevents any single system from reading the data it holds. Typically, the data are then further encrypted locally on each system. Reading the data requires the cooperation of all the systems involved in splitting the data, otherwise there is not enough information to retrieve the original data. Erasure coding, a forward error correction code, is another technique often used as a sub-component of data dispersion methods. In schemes that involve sharding the data in the dispersion process, erasure coding helps prevent data loss.

This approach is commonly used in data centres or scenarios involving distributed data storage, as opposed to selectively encrypting messages within a secure communication protocol. These techniques are most often used in secure cloud storage, often in conjunction with redundant array of independent disks (RAID) solutions to improve the data security and privacy of client data.

2.4.3 Privacy-preserving digital identity

Traditionally, most systems use a centralized approach for identification and authentication. In this model, the service provider, which offers one or more services to users, fully controls the authentication process and manages the associated risks. The authentication information for users is stored on the service provider's side, but each user is required to remember their password or security question to access the services. As the use of internet services has increased, this requirement has become cumbersome for users who need to remember multiple credentials. This situation also introduces numerous security issues, as users often reuse credentials across various services.

Federated identity In an attempt to resolve these issues, researchers developed the concept of a federated identity model that allows users to reuse credentials across organizational and system boundaries. In this model, a trusted identity provider manages users' digital identities and allows them to use their credentials to access resources in other domains without needing separate accounts for each service. For instance, customers in Canada's commercial banks could use their banks' credentials to log into government services such as Canada Revenue Agency or Immigration, Refugees and Citizenship Canada.

Although federated identity management offers several advantages in terms of user convenience and users' privacy, it raises certain privacy considerations:

- **Data sharing and leakage**: Federated identity systems often involve the sharing of user attributes and claims (e.g., name, email address) between identity providers and service providers. Excessive or unnecessary data sharing can lead to the leakage of sensitive information, potentially violating user privacy.
- **User tracking**: When users authenticate with a federated identity, their activities across different service providers might be recorded and correlated by identity providers, creating a more comprehensive profile of their online behaviour. This tracking raises concerns about how this information is used and whether it is disclosed to third parties. For example, this information could potentially be used for targeted advertising or other purposes without the user's consent.
- **Legal compliance**: Different jurisdictions have varying data protection and privacy laws. Federated identity systems that span multiple countries or regions need to ensure compliance with relevant regulations, which can be complex. As a result, a digital federated identity from a specific identity provider may be usable for only a limited range of services.

Self-sovereign identity In recent years, a new decentralized model to manage digital identities, called self-sovereign identity (SSI), has emerged. This approach empowers users with greater control over their personal information and how it is shared and used. SSI has the following characteristics that could enhance users' privacy:

- **User-centric**: SSI places the individual at the centre of the identity management process. Users have full control over their identity information and can decide when and with whom to share it.
- **Decentralized identifier**: A core concept in an SSI is the decentralized identifier, or DID. DIDs are unique, persistent and cryptographically verifiable identifiers that are not tied to any centralized registry or authority. DIDs are the foundation for building SSIs.

- **Verifiable credentials**: In an SSI, individuals can issue, receive and store verifiable credentials. These are digital attestations or claims about a person's identity, qualifications or attributes. Verifiable credentials are tamper-proof and can be easily shared and verified without the need for a central authority.
- **Selective disclosure**: With an SSI, individuals can selectively disclose only the specific pieces of information required for a particular transaction or interaction. This minimizes the exposure of sensitive data and enhances privacy.
- **Interoperability**: SSI standards and protocols are designed to enable interoperability across different platforms and ecosystems. This allows users to use their SSIs in various contexts, from financial services to health care to online shopping.
- **No central authority**: Unlike traditional identity systems, an SSI has no single central authority or gatekeeper. Instead, trust is established through decentralized consensus mechanisms and cryptographic proofs.

In this model, users have full control over their data. This model relies on users to create their own DIDs using public key cryptography and to have full control over them. Users can acquire verifiable credentials from trusted entities, such as governments, private companies or universities. Users manage DIDs, verified credentials and the associated cryptographic keys in their digital wallets (e.g., on smartphones). This model provides better privacy and security since no central parties have access to users' digital identities and the associated data.

In the realm of SSI, the current technological landscape shows significant momentum, in both industry and academia. With a surge of interest and investment, various government and private entities are exploring the potential of SSI. They are crafting standards to establish a robust framework for the seamless integration and interoperability of decentralized identity solutions. Despite this progress, challenges persist in the maturity of SSI technology. Researchers and practitioners are engaged in addressing complexities associated with scalability, security and user adoption, essential for the widespread implementation of SSI systems.

2.4.4 Federated learning

Federated learning (FL) was introduced by McMahan et al. (2016) as a method that enables the analysis and learning from data distributed across multiple owners without the need to share individual datasets. Unlike traditional data analytics, which process data as a single dataset, FL handles data in a distributed manner. A central server coordinates with participants, who aggregate only the essential data locally. Only the minimal information required for the learning task is sent to the central server, where it is integrated with similar data from other sources. **Figure 1** illustrates the operation of FL.

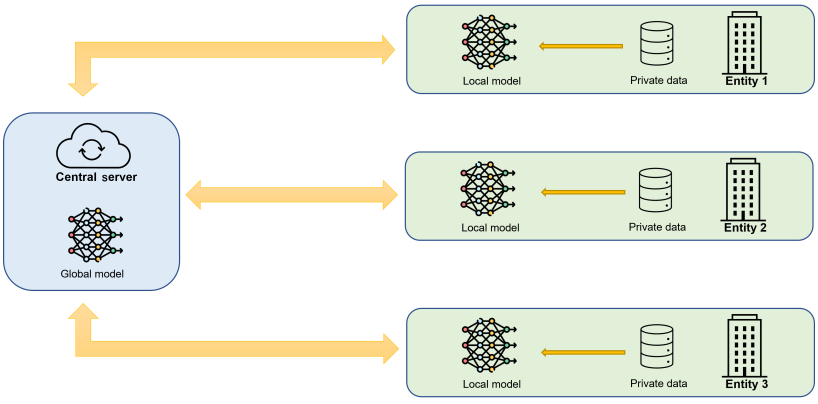

Figure 1: Federated learning

FL alleviates privacy concerns by ensuring that an individual's raw data is not transmitted to any external parties. It enhances user confidence by retaining only the pertinent data locally on the device. It treats the data on each device as separate training batches. Local updates to the global model are computed on each device and then sent to a central server for integration, depending on the learning algorithm.

Algorithm 1 Federated stochastic gradient descent

Initialize parameters θ
for each client i in N clients **do**
    Receive data D_i from client i
    Compute local gradient ∇_{θ_i}
    Send ∇_{θ_i} to server
end for
Aggregate gradients $\nabla_\theta = \sum_{i=1}^{N} \nabla_{\theta_i}$
Update parameters $\theta \leftarrow \theta - \eta \nabla_\theta$

However, FL faces challenges such as dealing with imbalanced and not independent and indentically distributed (non-i.i.d.) data. Additionally, concerns exist about the possibility of malicious participants or a central server compromising privacy. A malicious participant could potentially deduce the training data of others from the model updates sent to the central server, while the central server might infer sensitive information from the aggregated data (for more details on advances and open problems in FL, see (Kairouz et al. 2021)) .

2.4.5 Onion routing

The onion routing protocol in the Lightning Network is inspired by the Sphinx protocol (Danezis and Goldberg 2009), which is an extension of the onion routing technique used for anonymous communication over computer networks to provide privacy in transactions. Successive layers of hashed timelock contracts (HTLCs) and encrypted payment forwarding instructions are constructed off-chain to obfuscate the origin and destination of transactions. This ensures that, in cryptocurrency onion routing schemes, intermediaries do not necessarily know where they reside in the layered transaction, nor do they know the source and ultimate destination of the payment. Participants also are not aware of exactly how many other participants were involved in the payment layer.

The fundamental principle of onion routing is that the creator or originator of the onion message possesses sufficient information to establish shared secrets between the sender and each intermediary. This enables the encryption of the final onion payload in multiple, sequential layers. Additionally, as each layer of the onion is peeled away by intermediary nodes along the routing path, only partial routing information is revealed. This ensures that each intermediary node knows only the next destination to forward the onion message, without any knowledge of the overall path. Onion messages are also filled with pseudorandom padding to maintain a consistent fixed length, preventing observers or intermediary nodes from deducing their position in the routing path based on message length (Lightning Network RFC 2009).

3 Central bank digital currency design with privacy-enhancing technology

CBDC may involve the collection and management of PII (e.g., users' name, address, social security number and financial details including transaction history) to provide meaningful payment services to end customers, in observance with various laws and regulations to prevent illicit usage (e.g., *Proceeds of Crime (Money Laundering) and Terrorist Financing Act* (2000) in Canada), and to inform business decision-making. Due to the highly sensitive nature of these data, it is crucial to use privacy-by-design principles to incorporate privacy protections into the design and development of the CBDC system. As a result, while financial compliance (e.g., AML) is an important consideration for the CBDC system, these protections need to be balanced with privacy. The goal of this review is to explore the potential techniques to provide a high level of privacy to end users while ensuring that the system is compliant with the financial regulations and protected against illicit usage.

The design of CBDC will most likely involve trusted and authorized intermediaries between the central bank and end users to perform various operations, such as onboarding users. Privacy can be implemented through a combination of legal contracts for safeguarding PII, making it clear to entities what they can use data for, and through using technical controls such as encryption and access control. Our objective is to maximize the use of technology to implement customer privacy. Most of the current similar systems dealing with PII use conventional technologies such as access control, encryption for data at rest and in

transit, and consent management. These technologies have limitations in terms of unencrypted data during use, improper use of data when shared with external entities, unmanagable access control and ineffective consent management. Recently, with cryptocurrencies gaining the attention of cryptographers and computer scientists, advances in PET have shown promise in handling some of these challenges. We examine the use of these technologies in the design of a CBDC system and the challenges associated with it.

The design of a CBDC system consists of various modular components, each requiring careful attention to privacy and data protection to ensure overall privacy. We identify the following components as having a high impact on privacy due to their involvement in the collection, storage and processing of PII.

- **Onboarding service**: This service is partly responsible for verifying the information required to register a new user in the CBDC system. This process ensures due diligence related to regulatory, legal and credit requirements. It includes know-your-client (KYC) checks and procedures such as identity (ID) verification and document collection. This service may receive a (digital) identity from an identity provider and validate that the user presenting the ID is, in fact, the user the ID represents. It may also be a potential identity provider in the sense that it receives a collection of primary identification documents to construct a usable identity for purposes that meet the requirements of registering for CBDC. The profile constructed after onboarding will effectively act as the subset of PII that links transactions with actual users. Thus, the onboarding service must be carefully designed both in terms of technical implementation and governance to preserve privacy under non-criminal scenarios and be resistant to abuse.

- **Identity and access management**: Effective and secure authentication and access control are essential to prevent unauthorized access to sensitive data and to provide anonymous authentication to end users. This service applies mainly to various actors in the CBDC system, including end users and back-office administrators. For the end users, part of the onboarding process is to assign authentication credentials (e.g., username and password or biometric data). These need to be designed properly to ensure privacy for users when different entities involved in the CBDC system create spending profiles and manage sensitive information, especially biometric data of end users. Proper access control for back-office administrator users is also required to ensure that they have access only to the resources and functions they are authorized to use. Access control mechanisms, enhanced by PETs such as encryption and ZKPs, ensure that only authorized individuals can access sensitive data, in this way safeguarding privacy and ensuring regulatory compliance.

- **Transaction processor**: This component plays a crucial role in the processing, settling and recording of various types of transactions. A simple transaction consists of three main parts: *payer*, *payee* and *transaction amount*. However, it can contain various other fields (e.g., timestamp, device ID or payer location) known as transaction metadata, which can be used for purposes such as auditing, fraud prevention and analysis. Both main transaction data and metadata can raise significant risks if mishandled or misused. In addition to conventional technologies, PETs can potentially play an important role in the design of a transaction processor.

- **Compliance services**: The design of the CBDC system will include a service to ensure compliance with jurisdictional laws and regulations (e.g., *Proceeds of Crime (Money Laundering) and Terrorist Financing Act* (2000) in Canada) to prevent financial crimes. Typically, compliance requires that institutions with compliance obligations can view users' transactions (e.g., financial institutions) and have the authority to reveal suspicious activities to other entities (e.g., the Financial Transactions and Reports Analysis Centre of Canada, or FINTRAC [2]). Effective compliance also requires institutions to collaborate by sharing data. This raises many potential privacy issues where PETs can be used in the design to ensure effective compliance without unnecessarily revealing sensitive data.

- **Wallets**: Wallets are the key components of the system for individuals and entities to store, acess and transact CBDC. Wallets in the CBDC system will manage and have access to a great deal of PII. Along with providing security and functionality, the design must ensure the wallets protect the privacy of the individuals.

- **Data analytics**: CBDC data hold significant value for various purposes such as fraud prevention, monitoring and policy research. However, it is crucial that the design of the CBDC system safeguards individual privacy when internal and external stakeholders share this data. Certain PETs—such as

[2]FINTRAC is the financial transactions regulator in Canada.

altering data techniques, FL and SMPC—can be instrumental in meeting these objectives without violating the privacy principles of the system.

Although the development of PETs and their application to financial systems is experiencing great momentum, challenges still exist in the following areas:

- **Performance**: Can these techniques provide the desired scalability for a potential CBDC system (e.g., 10,000 transactions per second)?
- **Maturity**: Are there libraries that are tested and proven in the field?
- **Security**: Have these techniques gone through the rigour to clear the security analysis of a CBDC system?
- **Compliance services**: Can we perform required regulatory compliance (e.g., KYC and AML) when using these techniques since most of these techniques rely on hiding or masking the data?

Table 2 shows the application of PET techniques in the design of CBDC components.

Table 2: Summary of the use of privacy-enhancing technology techniques in central bank digital currency use-cases

Techniques	**Onboarding**	**IAM services**	**Transaction processing**	**Compliance services**	**Data analytics**	**Wallets**
Fully homomorphic encryption			✓	✓	✓	✓
Secure multiparty computation	✓	✓	✓	✓	✓	✓
Special signatures	✓	✓	✓			✓
Cryptographic commitment			✓			✓
Altering data techniques				✓	✓	
Self-sovereign identity	✓	✓				✓
Federated learning					✓	
Privacy-enhancing hardware		✓	✓		✓	✓
Zero-knowledge proof	✓	✓	✓	✓		
Onion ring			✓			
Transaction tumbler			✓			

4 Onboarding

Onboarding is the process of introducing a new user to a target system and is often tied to KYC protocols to comply with local AML and anti-terrorist financing (ATF) regulations. Efficient and cost-effective

onboarding is a critical component of the overall CBDC system. Digital identity credentials can play a major role in facilitating the onboarding process. If a user has already completed KYC with a trusted financial institution, the system could leverage these existing KYC credentials to onboard the user with another provider, eliminating the need to repeat the process. Successful onboarding would result in the issuance of valid authentication credentials for the user.

In the context of CBDC systems, onboarding users can be facilitated through both virtual and in-person methods, ensuring flexibility and accessibility. Virtual onboarding leverages digital platforms to verify users' identities remotely, using tools such as video calls, biometric verification and the submission of electronic documents. This approach offers convenience and efficiency, allowing users to complete the process from anywhere with access to internet and sufficient hardware. In contrast, in-person onboarding involves users physically visiting a designated location, such as a bank or service centre, where their identity can be verified face-to-face by a representative. This method may be preferred by those who are less comfortable with technology or who require additional assistance during the process. By offering both virtual and in-person onboarding options, the system can accommodate a diverse range of user preferences and needs, ensuring a more inclusive and user-friendly experience.

Processes of onboarding need to be balanced with financial inclusion and privacy objectives. For various non-criminal reasons, users may not have the necessary documents to sufficiently meet the standard of the service providers that are subject to identification requirements. Efforts are being made to enhance financial inclusion for underserved populations in banking and financial services while carefully avoiding the creation of new opportunities for criminal activity. These efforts can be facilitated with tiered onboarding with different levels of identification requirements. On one end of this tier, fully registered users can be onboarded by verifying the identification documents required by the law, and on the other end, non-registered users can be onboarded with no identification but have strict controls (e.g., total holding amount, transaction amount limits). We can potentially design other options with varying degrees of identification requirements and controls to satisfy different users' needs. The information collected during the onboarding process can be used to build a unique profile to assist with AML and other compliance investigations, or for data analytics purposes that are core to business operations such as risk management.

Offboarding is the process of removing or transitioning an existing user from a target system. This process is typically executed in accordance with established protocols and procedures to ensure compliance with regulatory requirements and security standards. Similar to onboarding, offboarding may also involve adherence to KYC protocols, especially in regulated industries, to facilitate the proper termination or transition of user accounts.

Since onboarding and offboarding are the first and last step of every observable transaction chain, this is where the engineering of privacy protections and data collection mechanisms is most crucial. In general, we follow two principles to balance privacy and compliance requirements: (1) minimize the number of entities that have visibility of PII; and (2) limit the data collected and PII to the minimum necessary to comply with privacy and regulatory requirements.

4.1 Privacy-enhancing technologies for the onboarding process

Onboarding users typically involves verifying user ID documents that are issued by other trusted entities, such as government agencies and regulated financial institutions. Individuals can use an SSI to maintain control over their digital identity records and selectively share it for the CBDC onboarding process (Soltani, Nguyen, and An 2018). If an SSI system exists within the jurisdiction, the CBDC system can provide significantly more privacy to end users. SSI enables users to store their digital identity information in a wallet on their personal devices, ensuring that they have full ownership and authority over their data. When KYC is conducted for onboarding, only relevant selective information from the SSI is shared for verification and validation purposes. This allows users to present specific credentials or attributes without revealing their full identity information, enhancing the level of privacy and security.

SSI solutions can leverage a decentralized trust network, such as blockchain or distributed ledger technology, to verify the correctness and integrity of the credentials and the authenticity of the issuers. The SSI-based verification process is not visible to the credential issuers, providing a higher level of privacy to the users. Cryptographic techniques like ZKPs can also be used to verify the authenticity of the identity and protect the confidentiality of sensitive information. Depending on the construction of the ZKPs, varying amounts of information may be revealed to the verifier. This can enable conditional linkages between SSI and transactions. Consequently, an SSI may be partially linkable or unlinkable based on the specific verification protocols used.

We see a potential for using this technology for the efficient and cost-effective onboarding of users to the CBDC system while enhancing security and privacy. However, our feasibility analysis of SSI for the CBDC system identifies several challenges:

- **Interoperability issues due to lack of standards**: A major challenge arises from the absence of uniform standards, resulting in interoperability difficulties among various identity providers.
- **Absence of sufficient wallet solutions**: A crucial aspect of implementing SSI involves a developed ecosystem of wallet solutions, a lack of which poses a hurdle in realizing a seamless integration of this identity framework.
- **Technology maturity concerns**: Another challenge is the relative immaturity of the technology, which has not yet evolved to adequately support the stringent requirements of mission-critical systems such as identity issuance/verification and CBDC systems. This necessitates a careful approach to implementation.
- **Limited adoption by public and private entities**: The slow adoption of SSI by both public and private sector entities constitutes a significant challenge, potentially hindering the widespread acceptance and effectiveness of this identity paradigm.
- **Availability of legal IDs through SSI systems**: Not all users may possess or have access to digital identification that meets the necessary verification standards.

5 Identity and access management services

To guarantee that operations involving CBDC accounts and digital wallets are restricted to authorized users, it is essential to properly implement authentication processes for users, devices and transactions. From an access control standpoint, we must ensure that users have access only to their own funds and transaction records.

Assuming that a user acquires credentials upon onboarding with a designated financial institution or entity, authentication is a critical procedure for associating these credentials with payment messages. This ensures that activities within the payment system occur legitimately and with the explicit consent of the original owner. For registered users, the system needs to frequently verify the identity of participants who are interacting with it and conducting transactions in order to uphold authenticity and deter fraudulent activities. Furthermore, for compliance with various regulations, the system must link authentication credentials to the owner's identity (PII), which emphasizes the importance of designing such processes to safeguard user privacy. While incorporating additional information or factors can aid authentication in fortifying system security, at the same time it amplifies concerns about user privacy.

Privacy concerns arise regarding how user information transmitted from devices like payment smartcards, mobile devices or personal computers during online transactions is processed, stored and managed by merchants or online commerce platforms. Consumers are also apprehensive about potential leaks and the linking of this information with other data. They are equally concerned about how financial institutions handle and safeguard their authentication information. Consequently, it is mandatory that all authentication data used in a transaction be deleted from the recipient's system unless the user gives explicit consent. Additionally, authentication data must be securely protected during transit and processing.

Various authentication methods are based on three distinct factors: *knowledge* (what you know), *possession* (what you have) and *identity* (who you are). Traditional authentication, such as username and password, relies on what the user knows. Methods such as security tokens, credentials, user devices or smartcards validate what the user possesses. Biometric authentication, using techniques such as face recognition or fingerprint scanning, verifies the user's identity. Multifactor authentication (MFA) enhances security by combining at least two of these factors. For instance, a common MFA implementation might entail entering a password (what you know), receiving a one-time passcode on a mobile device (what you have) and providing a fingerprint scan (who you are).

5.1 Privacy-enhancing technologies for identity and access management services

In this section, we offer high-level overviews of various PET techniques for privacy-preserving authentication. Generally, these methods can be categorized into two primary types: *linkable credentials* and *unlinkable credentials*. The linkability of a credential indicates whether a trusted entity can reconstruct a link between PII and transactions using only data internal to the system. However, this does not preclude the possibility of reconstructing the mapping through alternative or external datasets not involved in the system's operation. It is important to note that this possibility exists even with credentials that are unlinkable and anonymous.

According to this definition, a credential can be anonymous yet linkable if only a trusted entity can reconstruct the mapping while remaining effectively anonymous to other entities or participants in the system. Conversely, a credential cannot be both non-anonymous and unlinkable, and it is inherently linkable if it is not anonymous. Additionally, a credential can be conditionally linkable, where information is disclosed only under specific conditions, enforced by cryptographic techniques such as a ZKP.

5.1.1 Linkable credentials

In a linkable credential scheme, a client provides their PII to a specific organization, such as a financial institution. Once the organization verifies the client's PII, it issues credentials to the client, which can then be used to conduct transactions. It is important to note that the issuer never holds the client's private key in any of these scenarios. In linkable schemes, verifiers have a way to obtain more details about the identity associated with the credentials, which can be advantageous in jurisdictions with strict compliance demands while still restricting access to sensitive information.

- **Credential generation via dynamic group signature**: During onboarding the user receives a private key and an ID as part of the credential from a financial institution. The client ID is linked to an account (or may also serve as an account). A user signs their transaction using their private key. The fully dynamic group signature scheme (Delerablée and Pointcheval 2006) ensures that no two transactions can be linked by anyone except the trace manager. This level of privacy is classified as anonymous since, other than the trace manager, observers gain no information from observing the transaction.

- **Credential generation via digital signatures**: In these protocols, during onboarding a user receives a pair of public/private keys for a digital signature scheme. The financial institution associates the user's account with the public key and the user signs the transaction with their private key. The transactions are verified using the corresponding public key. It is important to note that anyone can link two transactions in this scenario. Thus, using digital signatures alone offers only pseudonymity.

5.1.2 Unlinkable credentials

A client can submit their PII to a designated organization for verification. Once the organization verifies and validates the client's information, it issues credentials to the client using PETs. An alternative method would be to allow the client to generate their own anonymous credentials, with the information behind these credentials verified and validated in zero-knowledge. Subsequently, the credentials can be used for transactions when authenticated by the owner. In these unlinkable schemes, linkages cannot be derived using only the internal information of the system, which means a greater effort is needed to meet compliance requirements. This may lead organizations to record additional metadata and linkable data outside the transactional system to comply, which could undermine the overall privacy protections the system was designed to uphold.

- **Credential generation via anonymous credentials**: A user and an issuing authority participate in a ZKP protocol, at the conclusion of which the user receives an anonymous credential. The organization then recognizes the user as part of the verified group. The user can subsequently register with any financial institution by presenting a ZKP of their membership. Several models of anonymous credentials (Camenisch and Lysyanskaya 2001; Camenisch and Lysyanskaya 2004; Baldimtsi and Lysyanskaya 2013) are suitable to implement in such an onboarding system.

- **Credential generation via multiparty computation**: In an alternative approach, multiple entities collaboratively assemble a credential, with each entity holding a piece of the credential. These pieces

are combined using an SMPC protocol. The credential can be reconstructed and verified with the cooperation of some or all participants in the protocol. However, without reaching a specific threshold, the credential lacks sufficient information to establish linkages or map the identity to its transactions.

Once a user obtains credentials confirming successful onboarding, they can use these to authenticate in a privacy-preserving manner before conducting operations in a CBDC system. In a CBDC system, successful authentication grants users access solely to their own holdings. However, this linkage of an authenticated user to their holdings and potentially other information could jeopardize the user's privacy.

Using different PETs, we can achieve varying levels of privacy, subject to compliance with regulatory laws. Authentication using linkable credentials may allow authorized agencies to connect accounts and transactions, whereas authentication with unlinkable credentials through anonymous credentials or SMPCs enhances user privacy. For example, by using an anonymous credential to authenticate in a CBDC system, a user leaves no trace that could allow an authorized agency to link their account to subsequent operations. The user anonymously demonstrates to a verifier that they possess a valid credential. An issuer can revoke a user's privileges easily. However, even if all parties (including other users, verifiers and the issuer) were to collude, they would not be able to determine the user's identity from the proof of validity (Baldimtsi and Lysyanskaya 2013).

Likewise, authentication using SMPC works when multiple validating parties collaborate to set up the SMPC protocol, which involves generating cryptographic keys, distributing shares of those keys to the participants and establishing the necessary communication channels. These parties can use an SMPC-based credential issuance protocol to provide the credential to the user. An authenticating user inputs the share of credentials (e.g., password, biometric data or cryptographic keys) to each of the validating parties. Validating parties engage in the secure computation protocol using SMPC to collectively compute a function that verifies the authentication credentials. One could design a protocol where both credentials and holdings are shared among SMPC participants so that no one party on its own knows the holding account that the user is operating in. With the careful design of the protocol and cryptographic techniques, it might be possible to achieve high throughput and low latency. However, the complexity and the maturity of the anonymous credential and SMPC protocols remain major challenges in implementing a solution using these techniques.

6 Transaction processing

A core aspect of a CBDC system involves managing the transfer of ownership of funds and efficiently processing these transactions on a large scale. This crucial service is responsible for handling all aspects of data storage as well as the operational logic for transaction processing, clearing and settlement across the entire CBDC network. This involves not only maintaining transaction integrity but also ensuring that transactions are processed swiftly to meet the demands of a modern economy. It should also be scalable to handle growth in transaction volume as adoption of the CBDC increases. Integrating advanced technologies such as machine learning could further enhance monitoring and predictive maintenance capabilities, ensuring the system's reliability and efficiency over time.

Additionally, this component must have the capability to track and report the status of any payment instruction that has been submitted. This is vital for providing transparency and accountability in financial operations. It should also be able to furnish detailed data regarding transactions and endpoints upon request, especially to the compliance component of the system. This information is essential for regulatory compliance, helping to ensure the system adheres to legal standards and best practices in financial operations.

To maintain provenance and integrity of funds, traditional digital transaction processing systems record and log many data artifacts to leave an audit trail. For consumer privacy, the following is required: (1) the system should collect only the transaction data needed to process payments and satisfy legal obligations; (2) transactions should not be linkable; and (3) transaction data should not be shared with outside parties, except law enforcement for legal reasons.

6.1 Privacy-enhancing technologies for transaction processing

Tokenization is increasingly used in the card payment industry to protect sensitive information, such as the primary account number on credit cards. This method substitutes the 16-digit credit card number with a unique, randomly generated token. These tokens are typically created dynamically for each transaction or when a card is registered with a system. Merchants and payment processors then use these tokens to initiate and process transactions instead of the actual credit card numbers, enhancing user anonymity during

payments. This strategy also reduces the risk of privacy breaches by limiting the exposure of sensitive information in the event of a security breach at merchants or service providers. A similar tokenization approach could be adopted for device-based CBDC systems, such as smartcards. By tokenizing unique device identifiers, consumers can securely and anonymously make online purchases and link their CBDC devices with their smartphones or other smart devices.

Various PET techniques have been introduced in digital currencies and cryptocurrencies. The first PET technique for an anonymous payment was blind signature, introduced by Chaum about 40 years ago. The purpose of Chaum's eCash is to make payments anonymous from the issuing bank—that is, the bank does not know who pays for what. In a Swiss National Bank working paper, Chaum, Grothoff, and Moser (2021) describe how to issue a CBDC based on Chaum's blind signatures. The GNU Taler project implemented a privacy-friendly payment system based on blind signatures (GNU Taler. n.d.). While these systems provide good privacy for consumers, they can only be used for online transactions. Furthermore, a blind signature-based payment system can issue fixed denomination coins (e.g., \$1, \$5, etc.), which raises an issue with change. Another challenge of using blind signatures is the lack of standardized protocols. The only protocol standardized so far is RSA blind signatures (Denis, Jacobs, and Wood 2023), which is not quantum-safe.

Recent development in blockchain technology has introduced a range of PET techniques that could be applied to provide privacy for senders, receivers and transactions. To protect the identity of a transaction's sender, one can use (one-time) ring signatures (see Section § 2.3.1). Monero (van Saberhagen 2013) was the first cryptocurrency that implemented a one-time ring signature into its blockchain. This signature scheme allows users to achieve unconditional unlinkability. The idea is to conceal the identity of the sender among a group of users by producing a signature that can be verified by a set of public keys rather than the sender's key. The sender is hence indistinguishable from the other users whose public keys are in the set.

The Monero blockchain uses the Pedersen commitment (see Section § 2.3.3) to conceal the amount of transactions. This technique was used in implementing a transaction scheme known as RingCT (Noether 2015), with the first version being published in January 2017. In this system, a commitment C for a transfer amount v uses a random value as a binding factor, making it appear as a random value across the network. The sender can later disclose both the binding factor and the amount, enabling others to verify the correctness of the commitment submitted. Although observers or verifiers on the Monero network cannot determine the amount of a specific transaction, they can verify that the sum of the input amounts is equal to the sum of the output amounts, thanks to the additively homomorphic property of the Pedersen commitment. Cryptographic commitments are implemented in conjunction with range proofs, which are cryptographic mechanisms that prove a value falls within a specific range without revealing the actual value. The Monero network uses range proofs to ensure that none of the outputs are negative.

The onion routing of the Lightning Network on Bitcoin can enhance the privacy of senders and receivers in transactions. In this protocol, intermediaries do not necessarily know where they reside in the layered transaction, nor do they know the source and ultimate destination of the payment. Participants also are not aware of exactly how many other participants were involved in the payment layer. However, the attacker can make some association or infer correlation of packets through a combination of traffic analysis and sybil attack [3] , by comparing the fundamental network topology of the Tor network with the Lightning Network. The Tor network's topology resembles a fully connected graph, whereas the Lightning Network's topology resembles a graph with weakly connected components made up of numerous star-shaped subgraphs (Romiti et al. 2021).

Blockchain technology uses a PET called stealth addresses to safeguard the privacy of recipients. This technique allows the sender to create a unique, one-time-use payment address for each transaction directed to a specific receiver. As a result, observers on the network are unable to link multiple transactions to the same receiver until those funds are used in a subsequent transaction. The receiver—by leveraging their private key, which corresponds to the stealth address, along with the ephemeral key—can derive the secret key for the one-time-use payment address that is necessary to access the funds. Stealth addresses are used in cryptocurrencies such as Monero and Zcash (Hopwood et al. 2021) to bolster the privacy of recipients.

ZKP (see Section § 2.3.4) is a powerful PET that preserves the privacy of both the sender and the transaction amount. While ZKPs have a long history in cryptography, their first practical application emerged with blockchain technology. Zcash was the pioneering blockchain to implement a ZKP protocol known as ZK-SNARKs (Ben-Sasson et al. 2018). This technology allows individuals to prove certain statements without revealing the data underpinning the proof, enabling a consumer to authenticate a transaction without disclosing any PII. Additionally, by using a ZKP protocol, a consumer can demonstrate

[3] A Sybil attack is a security threat where a single entity creates multiple fake identities to manipulate or disrupt a system.

sufficient funds for a transaction, hide the transaction amount and confirm these details with the verifiers.

Integrating ZKPs with a transaction tumbler not only enhances the confidentiality of the transaction amount but also hides the identities of both senders and receivers (Bitcoin Wiki n.d.). A notable example of this implementation is Tornado Cash, which allows users to deposit funds of fixed denominations into the protocol while simultaneously generating a secret for a UTXO. This secret enables the original sender to transfer funds to a recipient in zero-knowledge, as the system only verifies the possession of a valid secret without disclosing its associated UTXO. When the sender decides to execute the transfer, they can either (1) create a ZKP that includes proofs of transaction validity, recipient details, nullifier hash and fees for a relayer who then submits the transaction data to the on-chain smart contract on behalf of the sender; or (2) send the secret directly to the recipient who must then independently withdraw the funds.

Last but not least, SMPC is a PET that ensures the privacy of senders and receivers and conceals the transferred amount. A sender, authenticated by a financial institution with valid credentials, can initiate a transaction by splitting their credential, account number and transaction amount and then distributing these shares to multiple validating parties in the system. These parties engage in an SMPC protocol to first authenticate the sender (as outlined in Section 5.1). Subsequently, they collaboratively determine if the sender has sufficient funds for the transfer and if the amount complies with established limits (e.g., $10,000). In the final phase, the parties use another SMPC protocol to transfer the specified amount from the sender's account to the recipient's account. Throughout this process, no single party has access to any information about the transaction amount, sender, receiver or their account balances.

The integration of cryptographic PETs into a CBDC transaction processor presents both opportunities and challenges. While these technologies enhance privacy, they also involve complex mathematical computations that can introduce computational overhead, potentially affecting transaction processing speed and the scalability of the CBDC system. Another challenge is the maturity and complexity of these techniques, which may not yet be sufficiently developed to meet the robust demands of a mission-critical system such as CBDC. This immaturity can lead to reliability and performance issues that are critical in financial systems.

Interoperability with existing financial systems and infrastructure also poses significant challenges. Current systems use standardized payment message formats, such as ISO 20022 (n.d.), and we do not yet know how messages from these PETs will align with those of legacy financial systems. The lack of universally accepted standards for ZKP, SMPC, or specific signature implementations could complicate seamless integration across various systems.

Moreover, achieving compliance with regulations such as AML becomes more complex when using these PETs. For instance, while ZKPs can confirm that a transaction complies with certain criteria without revealing underlying data, they make it challenging for regulators to trace the origins of funds or identify patterns of money laundering. This can create significant hurdles in enforcing AML regulations and other compliance requirements.

Finally, user adoption and education pose significant challenges. Users may not be familiar with concepts like ZKPs or SMPC, underscoring the need for effective education and communication to build trust and facilitate acceptance.

7 Compliance services

In many jurisdictions around the world, regulations govern the permissible use of data and PII by any organization, whether it is a private company or a public institution. Particularly in financial or transactional contexts, additional, stringent regulations often mandate service providers to collect information to comply with AML, KYC requirements, and anti-fraud measures. Specifically in Canada, firms generally fall under the *Personal Information Protection and Electronic Documents Act* (2000), and the federal government has proposed new rules about consumer privacy and other protections in Bill C-27, which is currently before the House of Commons (Library of Parliament 2022). Additionally, financial institutions are also subject to the *Proceeds of Crime (Money Laundering) and Terrorist Financing Act* (2000), as well as regulations from FINTRAC, and adhere to guidelines from international regulatory bodies like the Financial Action Task Force.

In systems that serve as definitive records for asset ownership, significant incentives and opportunities exist for criminal and fraudulent activities. Mechanisms that ensure the integrity and correctness of such systems typically can guarantee these properties only within the confines of the system itself. Therefore, a compliance system is essential for monitoring adherence to the system's correctness and to other relevant legal

and regulatory frameworks concerning information that crosses the system's boundaries. This compliance system should combine human oversight and governance with automated processes and technical solutions designed to boost efficiency and identify both recurring and unusual malicious activities.

Typically, rules-based compliance systems are straightforward to automate because they operate on clearly defined rules and behaviours. For instance, any transaction exceeding $10,000 must be reported to the regulatory authority. Principles-based compliance frameworks consist of broad guidelines designed to direct behaviour across various scenarios, making them more challenging to implement in practice. These frameworks typically depend on human judgement and intervention. However, advancements in technology are increasingly supporting the efficiency of compliance processes through improved and earlier detection capabilities. When a behaviour pattern or transaction is flagged as potentially suspicious, system operators are often required to report and assist in investigations conducted by regulatory authorities. In more extensive investigations with stronger suspicion of criminal activity, examining only a few transactions may not yield sufficient detail to fully understand the activities of the suspected entity.

Entity resolution is a technique commonly used in middleware products that operate at the boundaries of systems to monitor and manage complex record linkages among various participating entities. This is often necessary because data from a single system may not provide a comprehensive understanding of the target entity. In environments with multiple intermediaries serving end users, the challenge is to collaborate to ensure efficient compliance through data sharing while simultaneously protecting proprietary business information and customer privacy. By pooling and linking data from various sources associated with the same entity, investigators can achieve a clearer insight into the activities of a suspicious entity, helping determine the presence of criminal activity or supporting effective enforcement actions against that entity. This technique is widely applicable and has been particularly effective in combating sanctions evasion and ATF.

7.1 Privacy-enhancing technologies for compliance services

Achieving a high level of privacy protection in the CBDC system requires minimizing the collection and storage of PII data. However, financial compliance requirements mandate the service providers to have access to user identity and transaction data and to share this information with other entities (e.g., for investigations). The conflict between privacy and compliance is inherent. Some PETs can play an important role in the design of compliance systems to protect CBDC against illicit usage while protecting the privacy of the end users.

In projects lacking a single service provider accountable for compliance, *viewing keys* have emerged as a potential solution, particularly in the public cryptocurrency domain, such as in the ZCash and Monero projects. These projects have implemented a feature that enables users to create a viewing key that grants an entity read access to specific transactions, fully disclosing the transaction details. Additionally, range proofs are frequently used in cryptocurrencies to verify the accuracy of transactions. CBDC systems can apply these techniques as well, allowing for verification that transactions comply with specific attributes without requiring direct access to the data. For instance, a system could be configured to ensure that transactions below a certain threshold remain private, with no entity having visibility into the transaction details

Entity resolution is particularly effective when firms are in partnership or belong to the same parent organization, as analysts typically have access to and permission to use each database system. However, collaboration becomes challenging when independent organizations are involved, as firms often have business and economic reasons to withhold their raw data. FHE, when combined with SMPC techniques, offers a potential solution in these cases by enabling organizations to perform analytical queries across different entities without exposing their raw customer data. Despite advancements in FHE design (e.g.,by Chillotti et al. (2020)), using FHE for real-time processing remains significantly challenging due to its performance limitations.

SMPC schemes and protocols, as described in previous sections, can be configured so that information can be decrypted and accessed when a certain threshold of collaborating entities is met. This enables the design of a compliance system that activates under specific conditions, such as the issuance of a warrant, allowing a court to compel designated participating entities to collaborate and disclose information about a suspicious transaction. This approach safeguards the general privacy of transactions from any individual system operator while still enabling detailed compliance activities.

However, employing SMPC in contexts such as CBDC systems presents significant challenges. The cryptographic constructs required for SMPC are complex and currently have very few prototype implementations, which complicates their deployment and scaling. Additionally, these cryptographic

constructs are inherently slow, which poses a substantial barrier to their use in environments where real-time or near-real-time transaction processing is crucial. We need to address these performance issues to make SMPC a viable option for widespread use in CBDC systems. Another significant challenge with implementing SMPC in CBDC systems is the scarcity of skilled professionals who understand the intricate cryptographic constructs involved. The complexity of these constructs requires a high level of expertise in both cryptography and system design, which is not widely available in the current workforce. This lack of specialized knowledge can hinder the development and secure implementation of SMPC protocols within CBDC platforms.

8 Data analytics

While it is essential to maintain the confidentiality of data, particularly PII, these data are invaluable for various operations within financial institutions. These operations include fraud detection, feature extraction, business modelling, financial computations and other forms of data analytics, such as publicly sharing trends related to CBDC use in specific geographic areas. Additionally, there might be requirements to share PII among CBDC service providers, which could include entities associated with central banks or collaborations between central banks and commercial banks, such as in efforts to detect money laundering activities.

Insensitivity in handling customer or transaction data can result in substantial costs for businesses due to penalties from regulatory bodies. Legal and compliance obligations, such as the General Data Protection Regulation (European Parliament and Council of the European Union 2016) and the Payment Card Industry Data Security Standard (PCI Security Standards Council n.d.), levy significant fines on organizations following a data breach. Furthermore, major data breaches can erode trust in the safety and security of systems, particularly those of national importance. Numerous regulatory bodies have been established to oversee how organizations manage or misuse private data.

8.1 Privacy-enhancing technologies for data analytics

A range of PETs are available that can be used to ensure data confidentiality during analysis. In this use case, the most commonly preferred category of PETs involves data alteration (see Section§ 2.2), which includes methods such as anonymization, pseudonymization, synthetic data and differential privacy). These techniques obfuscate data by processing them locally and modifying them—either by adding noise or removing identifying elements—before releasing them publicly, sending to a research laboratory or sharing with a third party. The goal of these methods is to transform customer data to reduce disclosure risk to an acceptable level while maintaining the data's utility, thereby ensuring that the quality of the published or shared data is preserved.

One of the challenges of these techniques is to ensure that information does not leak (risk of reidentification) through *linkage attacks*, which occurs when malicious users combine multiple sources of data. For example, a hacker can combine a shared data record containing "gender," "postal code" and "date of birth," with a public voter list that contains "name," "gender," "postal code" and "date of birth" to identify the customer.

Ensuring a balance between two properties—-data utility and disclosure risk—is another challenge in this class of techniques that must be resolved. Data utility refers to a measure of how useful a dataset is for a given task, and disclosure risk refers to the risk that a malicious user can use the protected dataset to derive confidential information on an individual. While altered data could provide a higher level of privacy by adding more noise, the data analysis itself may not return insightful results if the underlying quality of the data has deteriorated through this process.

The next category of cryptographic PETs comes from the ability to compute on encrypted data, allowing computations to run over data that are never visible or disclosed. In contrast to data altering techniques, here the underlying data remain unmodified but are hidden by encryption. The idea was first introduced by Rivest, Adleman, and Dertouzos (1978). Despite garnering little industrial attention early on, this technique is quickly becoming a practical reality due to the recent developments of FHE and modern computers. The first FHE, Gentry's scheme (Gentry 2009), was very inefficient due to its costly bootstrapping process. But a new form of FHE, called Torus-FHE (Chillotti et al. 2020), with efficient programmable bootstrapping has led to huge performance advantages. However, FHE still lacks the performance capabilities to make it viable for real-time processing.

Another technique is SMPC, which enables two (or more) parties to collaboratively perform analyses over their input data while keeping those input data private. SMPC can aggregate sensitive data without

requiring any parties to disclose their own data (see more in Section 2.4.1). SMPC allows institutions to conduct analysis on private data held by other institutions without ever revealing those inputs. A challenge in an SMPC system is to prevent cross-reference attacks on the output and other information that attempt to infer sensitive data. Any party could carry out this attack in the protocol if they learned the true and exact output and conducted statistical analysis to reconstruct a subset of the original dataset (Lindell 2003). To prevent this, one could apply obfuscation techniques such as differential privacy to the outputs of an SMPC system.

FL can also provide data confidentiality by allowing multiple parties to collaboratively train a machine learning model in which each party uses their own dataset—that is, without sharing data. A central server can be used to coordinate the parties. Although data are not shared among parties or with the central server, one challenge is the imbalanced and non-i.i.d. data partitioning across parties. Malicious participants could feed such data into an FL system to gain insights from the trained machine learning model (see Section 2.4.4 for more details).

Finally, performing data analysis in hardware-based TEEs or confidential computing environments can help enhance assurance in data privacy. With on-premises systems becoming less common, the demand for third-party compute resource providers to prove the security attributes of their systems has grown. A dedicated logical boundary in a system can hold and process sensitive data and can run secure code within its confines. It can attest to the provenance of the hardware from trusted sources and the security configurations of the environment and can demonstrate its tamper-resistance. Even if the operating system is corrupted, the data stored and processed inside a TEE could not be accessed or exposed. Major chip manufacturers such as ARM, Intel and Qualcomm have implemented TEEs in their devices. Cloud providers such as Microsoft Azure, Google Cloud and AWS all offer confidential computing services in TEEs.

Cryptographic PET techniques such as FHE and SMPC offer significant advancements in data privacy, potentially prompting changes in legal frameworks regarding encrypted data processing. However, these technologies face operational challenges. For instance, once data is encrypted for use in FHE or SMPC, it becomes inaccessible for essential quality checks such as cleaning and preprocessing, which are crucial for ensuring the accuracy of data analysis. Moreover, FHE is computationally intensive, while SMPC demands substantial data communication, resulting in high operational costs. In addition, the lack of standardization in these cryptographic techniques complicates integrating them into existing systems and adhering to regulatory standards. This absence of uniform protocols can hinder widespread adoption, particularly in sectors such as financial services that require robust compliance with data protection regulations. Addressing these challenges involves not only technological improvements to enhance efficiency and reduce costs but also collaborative efforts to establish standards and update legal frameworks to better accommodate these innovative technologies.

9 Wallets

A wallet enables users to interact with a CBDC-based payment system. Typically, a wallet allows users to view their balance, make payments and receive CBDC from other wallets. A digital wallet consists of a data storage and computing environment where the customer's digital credentials are stored and used. Regardless of the specific CBDC design, digital wallets generally come in two types: (1) *custodial wallets*; and (2) *non-custodial wallets*. In the former, account providers offer support in deployment and management of payment credentials (e.g., private keys); in the latter, customers must maintain the payment credential themselves.

In custodial wallets, account providers such as commercial banks have two main options for managing payment credentials: they can either deploy these credentials directly to the customers' devices or keep them stored at the bank. In the former scenario, customers can authenticate and authorize transactions independently, without any need for bank interaction. In the latter scenario, customers must authenticate with the bank before they can perform any transactions on their accounts. The privacy of customers in custodial wallets might be compromised if the bank does not implement effective PETs.

In non-custodial wallets, users have full control over their payment credentials, which are stored directly on their own devices rather than with a financial institution. This arrangement allows users to execute transactions autonomously, without needing to interact with a bank or other account provider for transaction authentication. While this setup enhances user privacy by keeping the payment credentials solely under the user's control, it also places the responsibility of securing these credentials on the users. Consequently, although non-custodial wallets offer greater control and privacy, they also increase the risk of loss or theft

of funds if the user's device is compromised or if they mismanage their private keys.

Digital wallets can be designed as either hardware- or software-based solutions. Hardware wallets, like USB dongles or smartcards, provide robust security by physically isolating the digital assets from potentially vulnerable devices such as a user's smartphone. However, their limited resources may restrict their applicability to certain use cases. Regardless of the format, it is essential for payment credentials within these wallets to be stored and managed in a secure environment to ensure safety and integrity.

9.1 Privacy-enhancing technologies for digital wallets

As previously mentioned, the initial PET technique needed to secure digital assets in a wallet involves creating a protected environment. This can include hardware-assisted TEEs such as ARM TrustZone or tamper-resistant elements such as smartcards. While these trusted computing devices are widely used in existing digital wallets, they need to be evaluated for security vulnerabilities against sophisticated attacks, including backdoors and side-channel analysis.

Implementing memory encryption and MFA can further safeguard the confidentiality of digital assets stored in wallets. Encryption helps to protect data at rest by concealing them if the memory is dumped or accessed without authorization. Meanwhile, MFA enhances security by preventing data leaks if the device is lost or stolen. The second factor in MFA could involve a biometric scan, a personal identification number or a one-time password sent to a separate device. These layers of security can ensure that even if one barrier is breached, additional safeguards are in place to protect sensitive information.

10 Conclusions

This paper tackles the critical challenge of enhancing privacy in CBDC systems by introducing a comprehensive range of PETs and exploring their recent developments and potential applications. Through a deep analysis of these techniques, we evaluate their feasibility and viability for mission-critical systems such as CBDCs.

While significant progress has been made in the design and implementation of PETs, we find that challenges persist regarding their performance, maturity, and security in the context of CBDCs. Despite these challenges, we recognize the potential of PETs to revolutionize privacy in digital systems like CBDC. We emphasize the importance of continuous monitoring of developments in these technologies by organizations that collect and manage PII. As PETs mature and demonstrate viability in real-world applications, they should be considered for integration into CBDC system design and development.

References

Agrawal, D. and C. C. Aggarwal. 2001. "On the design and quantification of privacy preserving data mining algorithms." In *PODS '01: Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems,* 247–255. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/375551.375602.

Agrawal, R. and R. Srikant. 2000. "Privacy preserving data mining." In *SIGMOD '00: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, 439–450. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/342009.335438.

Al-Azizy, D., D. Millard, I. Symeonidis, K. O'Hara and N. Shadbolt. 2016. "A Literature Survey and Classifications on Data Deanonymisation." In C. Lambrinoudakis and A. Gabillon (eds.), *Risks and Security of Internet and Systems: 10th International Conference, CRiSIS 2015, Mytilene, Lesbos Island, Greece, July 20–22, 2015*, 36–51. Cham, Switzerland: Springer International Publishing. DOI:10.1007/978-3-319-31811-0_3.

Arm. n.d. "TrustZone for Cortex-A—Arm®." https://www.arm.com/technologies/trustzone-for-cortex-a.

Asrow, K. and S. Samonas. 2021. "Privacy Enhancing Technologies: Categories, Use Cases, and Considerations." Fintech Edge Special Report. Federal Reserve Bank of San Francisco. https://www.frbsf.org/research-and-insights/publications/fintech-edge/2021/06/privacy-enhancing-technologies/.

Baldimtsi, F. and A. Lysyanskaya. 2013. "Anonymous Credentials Light." In *CCS '13: Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, 1087–1098. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/2508859.2516687.

Bank for International Settlements. 2023. "Project Tourbillon: Exploring Privacy, Security and Scalability for CBDCs." Technical report. https://www.bis.org/publ/othp80.pdf.

Bank of England. 2023. "The Digital Pound: Technology Working Paper." Technical report. https://www.bankofengland.co.uk/-/media/boe/files/paper/2023/the-digital-pound-technology-working-paper.pdf.

Ben-Sasson, E., I. Bentov, Y. Horesh and M. Riabzev. 2018. "Scalable, Transparent, and Post-Quantum Secure Computational Integrity." IACR Cryptology ePrint Archive Paper 2018/046. https://eprint.iacr.org/ 2018/046.

Bitcoin Wiki. n. d. "Coinjoin." https://en.bitcoin.it/wiki/CoinJoin.

Blum, M., P. Feldman and S. Micali. 1988. "Non-Interactive Zero-Knowledge and Its Applications." In *STOC '88: Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, 103–112. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/62212.62222.

Brakerski, Z., C. Gentry and V. Vaikuntanathan. 2012. "(Leveled) Fully homomorphic Encryption Without Bootstrapping." In *ITCS '12: Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 309–325. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/2090236.2090262.

Brassard, G., D. Chaum and C. Crépeau. 1988. "Minimum Disclosure Proofs of Knowledge." *Journal of Computer and System Science* 37 (2): 156–189. https://doi.org/10.1016/0022-0000(88)90005-0.

Camenisch, J. and A. Lysyanskaya. 2001. "An Efficient System for Non-Transferable Anonymous Credentials with Optional Anonymity Revocation." In B. Pfitzmann (ed.), *Advances in Cryptology—EUROCRYPT 2001: International Conference on the Theory and Application of Cryptographic Techniques Innsbruck, Austria, May 6–10, 2001, Proceedings*, 93–118. Heidelberg, Germany: Springer. https://link.springer.com/chapter/10.1007/3-540-44987-6_7.

Camenisch, J. and A. Lysyanskaya. 2004. "Signature Schemes and Anonymous Credentials from Bilinear Maps." In M. Franklin (ed.), *Advances in Cryptology—CRYPTO 2004: 24th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15–19, 2004, Proceedings*, 56–72. Heidelberg, Germany: Springer. https://doi.org/10.1007/978-3-540-28628-8_4.

Chaum, D. 1983. " Blind Signatures for Untraceable Payments." In D. Chaum, R. L. Rivest and A. T. Sherman (eds.), *Advances in Cryptology, Proceedings of Crypto 82*, 199–203. New York, New York: Springer. https://doi.org/10.1007/978-1-4757-0602-4_18.

Chaum, D., C. Grothoff and T. Moser. 2021. " How to Issue a Central Bank Digital Currency." Swiss National Bank Working Paper No. 2021-03. https://www.snb.ch/en/publications/research/working-papers/2021/working_paper_2021_03.

Chaum, D. and E. van Heyst. 1991. "Group Signatures." In D. W. Davies (ed.), *Advances in Cryptology—EUROCRYPT '91: Workshop on the Theory and Application of Cryptographic Techniques, Brighton, UK, April 8–11, 1991, Proceedings*, 257–265. Heidelberg, Germany: Springer. https://doi.org/10.1007/3-540-46416-6_22.

Chillotti, I., N. Gama, M. Georgieva and M. Izabachène. 2017. " Faster Packed Homomorphic Operations and Efficient Circuit Bootstrapping for TFHE." In T. Takagi and T. Peyrin (eds.), *Advances in Cryptology—ASIACRYPT 2017: 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3–7, 2017, Proceedings, Part I*, 377–408. Cham, Switzerland: Springer International Publishing. DOI:10.1007/978-3-319-70694-8_14.

Chillotti, I., N. Gama, M. Georgieva and M. Izabachène. 2020. "TFHE: Fast Fully Homomorphic Encryption over the Torus." *Journal of Cryptology* 33(2): 34–91. DOI:10.1007/s00145-019-09319-x.

Confidential Computing Consortium. 2022. " Confidential Computing: Hardware-Based Trusted Execution for Applications and Data." White paper. https://confidentialcomputing.io/wp-content/uploads/sites/10/2023/03/ CCC_outreach_whitepaper_updated_November_2022.pdf.

Dalenius, T. 1986. "Finding a Needle in a Haystack or Identifying Anonymous Census Records." *Journal of Official Statistics* 2(3): 329–336. https://www.statcan.gc.ca/en/conferences/symposium2016/program/14731-eng.pdf.

Danezis, G. and I. Goldberg. 2009. "Sphinx: A Compact and Provably Secure Mix Format." In *S. P. '09: Proceedings of the 2009 30th IEEE Symposium on Security and Privacy*, 269–282. https://doi.org/10.1109/SP.2009.1.

Delerablée, C. and D. Pointcheval. 2006. " Dynamic Fully Anonymous Short Group Signatures." In P. Q. Nguyen (ed.), *Progress in Cryptology—VIETCRYPT 2006: First International Conference on Cryptology in Vietnam, Hanoi, Vietnam, September 25-28, 2006, Revised Selected Papers*, 193–210. Heidelberg, Germany: Springer. https://doi.org/10.1007/11958239_13.

Denis, F., F. Jacobs and C. A. Wood. 2023. "RFC 9474: RSA Blind Signatures." Internet Research Task Force (Crypto Forum Research Group) memo. https://doi.org/10.17487/RFC9474.

Dwork, C., F. McSherry, K. Nissim and A. Smith. 2006. " Calibrating Noise to Sensitivity in Private Data Analysis." In S. Halevi and T. Rabin (eds.), *Theory of cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4–7, 2006, Proceedings*, 265–284. Heidelberg, Germany: Springer. https://doi.org/10.1007/11681878_14.

ElGamal, T. 1985. "A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms." *IEEE Transactions on Information Theory* 31(4), 469–472. https://doi.org/10.1109/TIT.1985.1057074.

European Parliament and Council of the European Union. 2016. "General Data Protection Regulation." Regulation (EU) 2016/679. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

Fiat, A. and A. Shamir. 1987. " How to Prove Yourself: Practical Solutions to Identification and Signature Problems." In A. M. Odlyzko (ed.), *Advances in Cryptology—CRYPTO '86: Proceedings*, 186–194. Heidelberg, Germany: Springer. https://doi.org/10.1007/3-540-47721-7_12.

Gabizon, A., Z. J. Williamson and O. Ciobotaru. 2019. " PLONK: Permutations over Lagrange-Bases for Oecumenical Noninteractive Arguments of Knowledge." Cryptology ePrint Archive Paper No. 2019/953. https://eprint.iacr.org/2019/953.

Garms, L. 2020. "Variants of Group Signatures and Their Applications." Doctoral thesis, London, United Kingdom: Royal Holloway, University of London. https://pure.royalholloway.ac.uk/ws/portalfiles/portal/38498511/2020garmslhphd.pdf.

Gentry, C. 2009. " Fully Homomorphic Encryption Using Ideal Lattices." In *STOC '09: Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, 169–178. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/1536414.1536440.

GNU Taler. n.d. "Taler." https://taler.net/en/index.html.

Goldreich, O., S. Micali and A. Wigderson. 1987. " How to Play ANY Mental Game." In A. V. Aho (ed.), *STOC87: 19th Annual ACM Conference on Theory of Computing*, 218–229. New York, New York: Association for Computer Machinery. https://doi.org/10.1145/28395.28420.

Goldwasser, S., S. Micali and C. Rackoff. 1989. " The Knowledge Complexity of Interactive Proof Systems." *SIAM Journal on computing* 18 (1): 186–208. https://doi.org/10.1137/0218012.

Groth, J. 2016. " On the Size of Pairing-Based Non-Interactive Arguments." In M. Fischlin and J.-S. Coron (eds.), *Advances in Cryptology—EUROCRYPT 2016: 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part II*, 305–326. Heidelberg, Germany: Springer. DOI: 10.1007/978-3-662-49896-5_11.

Hopwood, D., S. Bowe, T. Hornby and N. Wilcox. 2021. " Zcash Protocol Specification." Electric Coin Company Technical Report. https://zips.z.cash/protocol/sprout.pdf.

Hush Hush. n.d. "Data masking definition." https://mask-me.net/datamaskingwiki/wiki/26/data-masking-definition. Accessed on October 12, 2023.

Intel Corporation. n.d. "Intel® Software Guard Extensions (Intel® SGX)." https://www.intel.com/content/www/us/en/products/docs/accelerator-engines/software-guard-extensions.html.

ISO 20022. n. d. "Universal Financial Industry Message Scheme." https://www.iso20022.org/.

Kairouz, P., H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, H. Eichner, S. E. Rouayheb, D. Evans, J. Gardner, Z. Garrett, A. Gascón, B. Ghazi, P. B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konecný, A. Korolova, F. Koushanfar, S. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Özgür, R. Pagh, H. Qi, D. Ramage, R. Raskar, M. Raykova, D. Song, W. Song, S. U. Stich, Z. Sun, A. T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F. X. Yu, H. Yu and S. Zhao. 2021. " Advances and Open Problems in Federated Learning." *Foundations and Trends in Machine Learning* 14 (1–2): 1–210. http://dx.doi.org/10.1561/2200000083.

Kargupta, H., S. Datta, Q. Wang and K. Sivakumar. 2003. " On the Privacy Preserving Properties of Random Data Perturbation Techniques." In *Third IEEE International Conference on Data Mining*, 99–106. DOI:10.1109/ICDM.2003.1250908.

Kilian, J. 1992. "A Note on Efficient Zero-Knowledge Proofs and Arguments (Extended Abstract)." In *STOC '92: Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing*, 723–732. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/129712.129782.

Kocher, P. C. 1996. "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems." In *Advances in Cryptology—CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18–22, 1996, Proceedings*, 104–113. Springer. DOI: 10.1007/3-540-68697-5_9.

Lamport, L. 1979. " Constructing Digital Signatures from a One Way Function." SRI International Technical Report No. CSL-98. https://www.microsoft.com/en-us/research/uploads/prod/2016/12/Constructing-Digital-Signatures-from-a-One-Way-Function.pdf.

Library of Parliament. 2022. "Legislative Summary of Bill C-27: An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts." 44th Parliament, 1st Session. https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/LegislativeSummaries/441C27E.

Lightning Network RFC. 2009. " Bolt #4: Onion Routing Protocol." https://github.com/lightning/bolts/blob/master/04-onion-routing.md.

Lindell, Y. 2003. *Composition of Secure Multi-Party Protocols: A Comprehensive Study*. Heidelberg, Germany: Springer-Verlag. https://doi.org/10.1007/b13246.

Ling, S., K. Nguyen, H. Wang and Y. Xu. 2017. "Lattice-based group signatures: Achieving full dynamicity with ease." In D. Gollmann, A. Miyaji and H. Kikuchi (eds.), *Applied Cryptography and Network Security—15th International Conference, ACNS 2017, Proceedings*, 293–312. Springer International Publishing. DOI:10.1007/978-3-319-61204-1_15.

Linux Foundation. 2003. "Confidential Computing Consortium." https://confidentialcomputing.io/.

Liu, K., H. Kargupta and J. Ryan. 2006. "Random Projection-Based Multiplicative Data Perturbation for Privacy Preserving Distributed Data Mining." *IEEE Transactions on Knowledge and Data Engineering* 18 (1): 92–106. https://doi.org/10.1109/TKDE.2006.14.

McMahan, H. B., E. Moore, D. Ramage, S. Hampson and B. Aguerra y Arcas. 2016. " Communication-Efficient Learning of Deep Networks from Decentralized Data." In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics* 54: 1273–1282. https://doi.org/10.48550/arXiv.1602.05629.

Mikhalev, I., K. Burchardi, I. Struchkov, B. Song and J. Gross. 2021. "CBDC Tracker." Oxford: Oxford University Press. https://cbdctracker.org.

Montjoye, Y.-A., A. Bourka, G. D' Acquisto, J. Domingo-Ferrer, P. Kikiras and V. Torra. 2015. " Privacy by Design in Big Data: An Overview of Privacy Enhancing Technologies in the Era of Big Data Analytics." Report by the European Union Agency for Cybersecurity. https://doi.org/10.2824/641480.

Noether, S. 2015. "Ring Signature Confidential Transactions for Monero." Cryptology ePrint Archive Paper No. 2015/1098. https://eprint.iacr.org/2015/1098.

Office of the Privacy Commissioner of Canada. 2017. " Privacy Enhancing Technologies—A Review of Tools and Techniques." Report prepared by the Technology Analysis Division of the Office of the Privacy Commissioner of Canada. November. https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2017/pet_201711/.

Pedersen, T. P. 1992. "Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing." In J. Feigenbaum (ed.), *Advances in Cryptology—CRYPTO 1991* 576: 129–140. Heidelberg, Germany: Springer.

PCI Security Standards Council. n.d. "PCI Data Security Standard (PCI DSS)." https://www.pcisecuritystandards.org/standards/pci-dss/.

Personal Information Protection and Electronic Documents Act, S. C. 2000, c. 5. https://laws-lois.justice.gc.ca/eng/acts/p-8.6/.

Pertsev, A., R. Semenov and R. Storm. 2019. "Tornado Cash Privacy Solution Version 1.4." University of California at Berkeley White Paper. https:// berkeley-defi.github.io/assets/material/Tornado%20Cash%20Whitepaper.pdf.

Polygon Labs. 2024. "zkEVM." https://polygon.technology/polygon-zkevm.

Proceeds of Crime (Money Laundering) and Terrorist Financing Act, S.C. 2000, c. 17. https://lois-laws.justice.gc.ca/eng/acts/p-24.501/.

Rivest, R. L., L. Adleman and M. L. Dertuzos. 1978. "On Data Banks and Privacy Homomorphisms." In R. A. DeMillo (ed.), *Foundations of Secure Computation,* 169–179. New York, New York: Academic Press, Inc. https://cdn.sanity.io/files/r000fwn3/production/c365f01d330b2211e74069120e88cff37eacbcf5.pdf.

Rivest, R. L., A. Shamir and L. Adleman. 1978. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems." *Communications of the ACM* 21(2), 120–126. https://doi.org/10.1145/359340.359342.

Rivest, R. L., A. Shamir and Y. Tauman. 2001. "How to Leak a Secret." In C. Boyd (ed.), *Advances in Cryptology—ASIACRYPT 2001*, 552–565. Heidelberg, Germany: Springer. https://doi.org/10.1007/3-540-45682-1_32.

Romiti, M., F. Victor, P. Moreno-Sanchez, P. S. Nordholt, B. Haslhofer and M. Maffei. 2021. "Cross-Layer Deanonymization Methods in the Lightning Protocol." In N. Borisov and C. Diaz (eds.), *Financial Cryptography and Data Security*, 187–204. Heidelberg, Germany: Springer. https://doi.org/10.48550/arXiv.2007.00764.

Samarati, P, and L. Sweeney. 1998a. "Finding a Needle in a Haystack." Harvard Data Privacy Lab.

Samarati, P. and L. Sweeney. 1998b. "Generalizing Data to Provide Anonymity when Disclosing Information." In *PODS '98: Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, 188. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/275487.275508.

Seničar, V., B. Jerman-Blažič and T. Klobučar. 2003. "Privacy-Enhancing Technologies: Approaches and Development." *Computer Standards & Interfaces* 25 (2): 147–158. https://doi.org/10.1016/S0920-5489(03)00003-5.

Shamir, A. 1979. "How to Share a Secret." *Communications of the ACM* 22 (11): 612–613. https://doi.org/10.1145/359168.359176.

Scoping SIG, Tokenization Taskforce. 2011. "Information Supplement: PCI DSS Tokenization Guidelines." PCI Security Standards Council. https://listings.pcisecuritystandards.org/documents/Tokenization_Guidelines_Info_Supplement.pdf.

Soltani, R., U. T. Nguyen and A. An. 2018. "A New Approach to Client Onboarding Using Self-Sovereign Identity and Distributed Ledger." In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 1129–1136. DOI: 10.1109/Cybermatics_2018.2018.00205.

Stapleton, J. and R. S. Poore. 2011. "Tokenization and Other Methods of Security for Cardholder Data." *Information Security Journal: A Global Perspective* 20(2): 91–99. DOI: 10.1080/19393555.2011.560923.

Saberhagen, N. V. 2013. "CryptoNote v 2.0." White paper, Monero Research Lab. https://www.getmonero.org/resources/research-lab/pubs/cryptonote-whitepaper.pdf.Xie, T., J. Zhang, Z. Cheng, F. Zhang, Y. Zhang, Y. Jia, D. Boneh and D. Song. 2022. "zkBridge: Trustless Cross-Chain Bridges Made Practical." In *CCS '22: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 3003–3017. New York, New York: Association for Computing Machinery. https://doi.org/10.1145/3548606.3560652.

Yao, A. C. 1982. "Protocols for Secure Computations." In *SFCS '82: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, 160–164. Washington, D.C.: IEEE Computer Society. DOI: 10.1109/SFCS.1982.38.

Zengo Ltd. n.d. "Zengo Wallet: Secure by Default." https://zengo.com.