Federal Reserve Bank of Cleveland Working Paper Series

Introducing a Framework for Measuring the Quantitative Benefits of Privacy-Enhancing Technologies

Ken Isaacson

Working Paper No. 24-16

August 2024

Suggested citation: Isaacson, Ken. 2024. "Introducing a Framework for Measuring the Quantitative Benefits of Privacy-Enhancing Technologies." Working Paper No. 24-16. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202416.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Introducing a Framework for Measuring the Quantitative Benefits of Privacy-Enhancing Technologies

Ken Isaacson

July 25, 2024

Abstract

This paper reviews privacy-enhancing technologies (PETs) and explores their benefits when used to make traditional payment processes more private. PETs can decrease privacy risk by reducing the amount of sensitive information accessible to payment-processing personnel and systems. This paper proposes a framework for quantifying the risk-reduction benefits of PETs. This method can be used to calculate the amount of privacy-risk exposure that may be created by a set of payment activities, estimate the amount by which PETs can decrease that exposure, and compare that quantified benefit against possible PET drawbacks. Assessing these drawbacks is outside the scope of this paper.

Keywords: Privacy-enhancing technologies, PETs, data, privacy risk, payments, personally identifiable information

JEL Codes: E42, O43

Ken Isaacson is a payments advisor and payments strategist at the Federal Reserve Bank of Cleveland and has broad experience in payments policy and operations, spanning revenue collections for US government agencies; the operation of the Fedwire Funds, Fedwire Securities, and National Settlement Services; and industry leadership to improve the US payment system by making it faster, safer, and more efficient. Ken received an MBA in finance from Columbia Business School and an AB in economics from Washington University in St. Louis (ken.isaacson@clev.frb.org). The author would like to thank Terri Bialowas, Susan Black, Paola Boel, and Ed Knotek for their helpful comments.

Introduction

Retail payment transactions often rely on payment messages that are transmitted between the payor and the payee through a set of intermediaries that may include financial institutions, payment-market infrastructure providers, and service providers. These payment messages contain information necessary for each party to process, settle, and apply the payment. Depending on the party, information required to perform its specific function may include sensitive personal information about the payor or payee, such as name, address, phone number, email address, identification number, or bank account number, among other details about the parties and the purpose of the payment. In the United States, the Gramm-Leach-Bliley Act (GLBA) states that personally identifiable information that a financial institution collects about an individual in connection with providing a financial product or service is generally considered "nonpublic personal information" subject to regulatory privacy protections.[1] In Europe, the General Data Protection Regulation (GDPR) protects individuals' personal data, which are defined as "any information relating to an identified or identifiable natural person" through an identifier such as name, identification number, location data, and so on.[2]

Consumers are concerned about the privacy of information they share. The International Association of Privacy Professionals found in a recent survey that 68 percent of consumers globally are somewhat or very concerned about their privacy online.[3] Indeed, consumer trust rises significantly when a company refrains from collecting information that is not relevant to its product.[4]

There are already a variety of privacy protections in place at financial institutions and other entities involved in the payment process. Nevertheless, there are periodic privacy incidents in which sensitive personal information is accessed by unauthorized individuals or is otherwise misused. In 2023, the number of data compromises in the United States increased to 3,205, up nearly 80 percent from those in the prior year and over 400 percent from those in 2015, affecting more than 350 million individuals last year.[5]

Data compromises are expensive, generating costs in multiple ways:

1. Operational costs: There is a patchwork of laws and regulations around the globe and within individual countries, sometimes specific to an individual region or industry, that define sensitive personal information, establish what constitutes a breach of that information, direct certain entities to notify consumers or authorities about possible breaches within specific time frames using specific notification methods, and establish penalties for failing to comply with associated requirements. In response, many companies and government entities establish privacy-incident response protocols that are complex and operationally burdensome to administer.[6]
2. Financial costs: Although not necessarily required by law, entities sometimes choose to purchase credit monitoring services for customers whose sensitive personal information has been breached. Credit monitoring services can cost between $0.25 and $2 per individual for one year of monitoring "depending upon the number of individuals, the type of information breached, and the services offered."[7]
3. Reputational costs: In a 2023 survey conducted by YouGov, more than 60 percent of British respondents would not trust a business with their personal data that has had a prior data breach. Ten percent of respondents further indicated that an organization cannot regain their trust after a data breach, regardless of actions taken by that organization in response to the breach.[8] The Motley Fool Ascent's 2024 Digital Banking Trend and Consumer Priorities Survey

additionally found that security and fraud protection features were among the most important factors for consumers in determining which bank they would select to open an account, while "good brand reputation" was also a top ten factor.[9]

4. End-user costs: Consumers and businesses may incur additional costs following a data breach related to fraud, embarrassment, or the need for additional security, for example. Such costs can vary based on the nature of the breached data and whether they were exposed, misused, or both.

Privacy-enhancing technologies (PETs) offer the possibility to reduce these costs by lowering the potential for privacy incidents. PETs comprise a suite of tools to keep sensitive information private when transacting payments and performing related functions surrounding the payment process. They are a collection of digital technologies and approaches that permit collection, processing, analysis, and sharing of information while protecting the confidentiality of personal data.[10]

In the recent industry literature, there is a significant focus on exploring PETs for use in emerging payment types such as digital currencies. The Bank of England points to PETs in its digital pound consultation as a possible way to allow payment providers to comply with legal and regulatory requirements without giving the government or the central bank access to users' personal data.[11] The Digital Dollar Project, a nonprofit forum exploring digital innovation in US-dollar-based money, considers PETs a potential way to support its aspiration for a central bank digital currency.[12] A Federal Reserve Board of Governors online discussion series featured a staff note that explored approaches to determining how, when, and where PETs should be used for digital asset payment systems.[13]

The benefits of PETs would apply to traditional payment processes, such as card payments, funds transfers, and automated clearinghouse payments, as well. Traditional payment processes sometimes involve storing, transmitting, analyzing, and processing sensitive personal information. This article will examine some of PETs' high-level benefits and drawbacks and propose a framework for evaluating their potential benefits when applied to traditional payment processes.

Privacy-Enhancing Technology Techniques

While there are several ways to categorize the different types of PETs, the Federal Reserve Bank of San Francisco offers a useful framework, dividing them into those that alter data, those that shield data, and those that use new systems and data architectures to make data more private.[14] Following Figure 1, select PET techniques will be described in more detail.

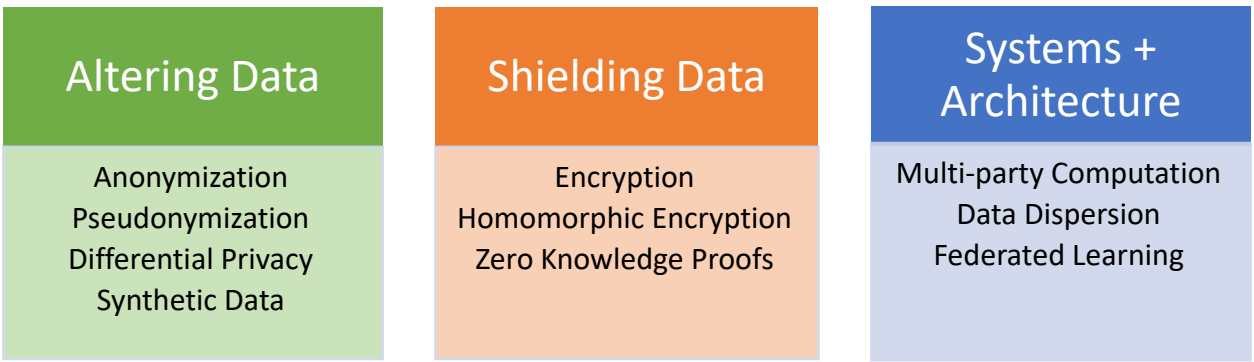

Figure 1: Categories of privacy-enhancing technologies (PETs) – modified from Asrow and Samonas, 2021[14]

Altering Data

The value of altering data stems from the concept that each party needs certain data in certain circumstances but not all data in all circumstances. Each data alteration technique allows for different levels of information to be transmitted, depending upon need. A marketing analyst at a retailer, for example, might benefit from knowing payment transaction patterns by time of day, time of year, product category, and store location, but they may not need to know the identities of the individuals making these payments. In such a case, the retailer may choose to pseudonymize the data by replacing each customer's name with a proxy name for that individual. This would allow for analysis of a given individual's unaltered payment patterns without knowing the individual's actual identity. If the analyst is more interested in aggregate payment patterns across customers, the retailer might choose to anonymize the data by removing customers' names altogether from the data set.

As another example, a quality assurance tester at a payment software company may need to test the software's ability to process a certain number of transactions per second carrying a certain amount of data per transaction and with a certain degree of field variation across different transactions. Using fabricated, or *synthetic,* data, might suffice in this case, with customers' names, addresses, and payment amounts randomly generated within defined parameters representative of real payment data that the system has historically processed. Alternatively, the tester may use differential privacy techniques that insert "noise," or additional arbitrary information, in targeted places that obscure sensitive information in the data set but still allow for testing of the software's proper generation of outputs when provided with specified inputs.

Shielding Data

Shielding data (i.e., protecting data from view) can be useful when there are parties processing data who can carry out their function without seeing the contents of the data. Encryption is a shielding method that uses mathematical techniques to transform data so that their original meaning is concealed. Payment data can be encrypted prior to being transmitted between different parties or when stored, making the data unreadable while in transit or at rest. When the data need to be accessed, they can be decrypted by a system or individual with access to the private encryption key, making them readable once again. A network provider or database operator, for example, would not be able to read encrypted payment data while in transit or at rest but could still perform its required function, moving the encrypted payment information from one point to another across the network or storing it in a database.

Homomorphic encryption can be used in circumstances in which an entity needs to make a calculation but does not need to see the underlying data.[15] A financial institution, for example, may have a unit that is responsible for checking that a payor has a sufficient balance in their account to fund a pending payment transaction. This balance-checking unit may be able to calculate the balance impact of the prospective payment using encrypted payment data without ever having access to readable payment details.

Zero knowledge proofs are yet another technique to conceal sensitive personal data. They could allow a party to prove that something about their credential is said to be true by a credentialling authority without having to share their full credential to the verifier.[15] Using this technique, a consumer, for example, could prove to an online liquor seller that the state has determined the consumer to be of legal drinking age without the consumer having to provide any specific state identification card details to the online seller. This would allow the seller to satisfy its legal obligation to verify that the consumer meets the legal drinking age without having access to the consumer's birth date, home address, identification number, or other sensitive information that may be contained on the consumer's credential.

Systems and Architecture

Systems and architecture can also be designed to break down sensitive data into smaller parts that are not sensitive when viewed in isolation. Multiple parties can then act in concert to perform a function such that each party is provided with access only to the smaller, less sensitive parts. Multi-party computation is an example of this technique, allowing each party to contribute to the collective calculation without giving any single party access to information that would be considered sensitive when paired with other pieces of information. Federated learning is another example in which insights can be centrally derived from data that are dispersed across multiple parties; the insights can then be shared without revealing the underlying sensitive data. As an example, it may be possible for a shared payment-fraud utility to look for patterns of bad actors that could be detected across data sets maintained by participating financial institutions without ever consolidating the data in one place. A participating financial institution would potentially be able to know when an entity identified in a customer's payment messages was frequently involved in fraudulent transactions reported by other institutions, all without having access to payment details at those institutions.[15]

The Payment Process

As is evident from some of the aforementioned examples, PETs hold promise in keeping data private while performing many functions surrounding the payment process. The end-to-end payment process can be complex and involve many entities; the specific entities involved vary by payment type and service provider.

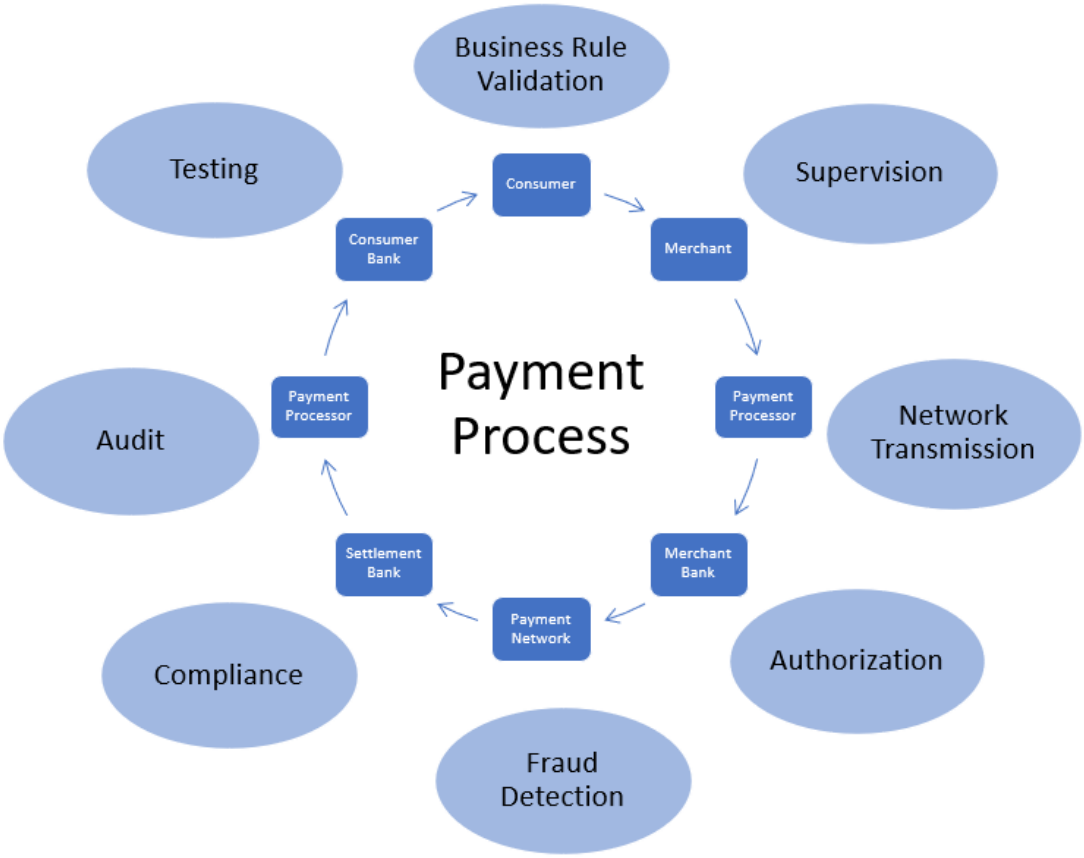

Figure 2: Example of an end-to-end payment process

Figure 2 portrays one way a consumer-to-merchant payment message could flow. The inner circle identifies the consumer making the payment, the merchant being paid, and each financial institution and service provider that could play a role in processing the payment. There are many other possible examples with different or additional parties that could play a role depending on the payment use case and the decision by each entity to insource or outsource specific aspects of its processing.

The outer circle identifies functions that typically surround a payment. Each function may involve one or more information systems performing a related action and one or more individuals with related business, technology, or operational responsibilities. It is common for each information system to duplicate data for resiliency purposes or to support testing. It is also common for multiple individuals to be involved in each function to support processing of the payment and to provide backup support to colleagues. These factors typically cause the number of copies of the data and individuals with data access to multiply.

Each entity, system, and individual involved in the payment process will need certain information to fulfill their duties. For example, payment authorization resources may need information only about the identity of the payor to authorize the payment. Business rule validation resources, for example, may need to determine that the number of characters in a payment field is within a specified limit; however, they may not need to know the content of the data in the field. Risk-management resources may need to

know the payment amounts and payor balances, but they may not need to know the purpose of the payment. If more information than needed is stored, processed, duplicated, analyzed, reported, or accessed, this could increase the likelihood that sensitive data will be compromised. Across the multitude of entities, systems, individuals, and business processes surrounding an end-to-end payment, this risk could be substantial.

Benefits of PETs

Many benefits of PETs are qualitative. PETs encompass techniques that can reduce exposure of sensitive information to individuals and systems performing payment-related functions while still allowing satisfaction of all requirements surrounding the payment process. PETs also may enable new value-added activities and forms of collaboration that were not previously feasible because of privacy concerns. Additionally, PETs could be used to build stronger, more effective models for analysis without the need to share underlying data to train these tools.[16] Processing payment transactions more privately could increase customer satisfaction, enhance brand image, and build customer trust.

To supplement an analysis of the qualitative benefits of PETs adoption, a quantitative methodology for benefit measurement is introduced below, focused on the value of reduced risk exposure made possible by PETs. The conceptual underpinning for the methodology is that there are a greater number of sensitive data fields shared with more organizations and accessible to more individuals at these organizations than would be necessary if the most appropriate PET for a particular circumstance were used. Additional sensitive data transmitted, stored, or accessed beyond what is needed to process a payment and to complete all related functions increases the surface area of the risk. The greater the surface area of the risk, the more likely it is that there will be a data breach and a subsequent cost incurred because of that data breach.

To quantitatively measure the privacy-risk-reduction benefits of PETs, the following definitions are offered:

Evaluator: The business unit or individual seeking to quantify the benefits of PETs adoption.

Boundary: All organizations, business units, supporting information systems, and supporting personnel with access to sensitive data to process payment transactions that are the subject of the PETs implementation evaluation. The boundary should be defined to include only those systems, entities, and business processes for which the evaluator has authority to make a PET implementation decision.

Sensitive Data Elements: The data elements present in payment messages and surrounding processes that the evaluator determines could cause harm if exposed to unauthorized individuals or if misused.

Point of Exposure (e): Each instance during which a person or information system involved in payment processing within the boundary has access to a single sensitive data element. Access to these data elements could be through a digital or physical record.

Total Exposure (E): The sum of all points of exposure within the boundary across all payment transactions processed over a selected historical comparison period (for example, the prior year) multiplied by the average number of copies of each sensitive data element maintained by each person or information system in physical or digital form. The process of calculating e and E will position the evaluator to more effectively identify potential points of excess exposure as per the definitions below.

Point of Excess Exposure (x): A point of exposure that is created through access to an unneeded sensitive data element. A data element is unneeded if its absence does not interfere with the information system's or person's ability to fully and effectively perform their duties. PETs have the potential to reduce or eliminate these excess points of exposure.

Total Excess Exposure (X): The sum of all points of excess exposure (x) over the selected historical comparison period, counting only points of excess exposure that could be eliminated by the specific PET being considered for implementation.

Probability of Breach (P): The evaluator's estimate of the probability of a breach generated by a single point of exposure over a one-year period. For illustrative purposes, industry data from 2022 indicate that there were 0.00014 records breached, on average, for each core noncash payment in the United States (Comparitech estimated 29.3 million financial industry records were breached in 2022 compared to the Federal Reserve Payments Study's estimate of 204.5 billion core noncash payments in the United States as of 2021).[16] Although this is not necessarily representative of the annual probability of a breach for any single point of exposure, it could be used as a reference point for the evaluator to baseline and adjust based on local circumstances.[17,18] To apply the referenced industry data on breach rate of records to this metric focused on breach rate of points of exposure, the evaluator may need to make several adjustments. For example, they may need to account for the number of records that exist per payment transaction, the number of points of exposure that exist per record, and potentially other factors relevant to the transactions within the boundary that would increase or decrease the breach rate compared to the industry data.

Cost Given Breach per Point of Exposure (c): The evaluator's expected cost in the event of a breach for each point of exposure affected by the breach. This could include financial-equivalent costs related to credit-monitoring services, forensic costs, legal support, hotline support costs, discounted future product and services costs, operational costs, loss of customer costs, and reputational costs, among others, that would be incurred because of the breach. For illustrative purposes, IBM estimates the average cost of a data breach in 2023 at $4.45 million, or $165 per record.[19] This could serve as a starting point for the evaluator to adjust up or down based on factors specific to the business activities and norms that apply to the payment transactions within the boundary. The evaluator may also need to convert the cost per record to a cost per point of exposure based on the evaluator's estimate of the number of points of exposure that exist per record.

Expected Breach Cost per Point of Exposure (C): The calculation of the cost of a breach for each point of exposure adjusted for the probability that there will be a breach in a given year.

*expected breach cost per point of exposure (*C) = *probability of breach* (P) * *cost given breach per point of exposure (*c)

By applying this proposed framework, the evaluator could calculate the financial-equivalent benefit of using PETs (B) as follows:

benefits (B) = *total excess exposure* (X) * *expected breach cost per point of exposure* (C)

The evaluator likely will want to consider these quantitative benefits, the key focus of this article, alongside other costs and benefits of PETs, evaluated separately from this framework.

Drawbacks of PETs

Payment and privacy professionals at corporations, financial institutions, market infrastructure providers, and service providers can evaluate the benefits of PETs relative to their drawbacks to help inform an assessment of the viability of adopting PETs. Several PETs' drawbacks are mentioned below and can be evaluated and quantified separately to complement the framework proposed above.

PETs may introduce financial costs, regulatory uncertainty, compliance challenges, and technology hurdles. Respondents to a recent Bank of England consultation paper found, for example, that PETs could introduce performance, security, and computational load trade-offs and introduce complexities that might reduce the ability to expand a system's overall scope.[20] PETs are also complex to implement, computationally intensive, costly, environmentally taxing, and difficult to govern given their complexity.[21]

Practical Steps

The proposed quantitative framework to measure the benefits of implementing PETs is a conceptual one that will have limitations in the real world. It may not be practical to obtain all model inputs with precise accuracy. There are nonetheless rough-order-of-magnitude risk-reduction insights that can be inferred from the benefits calculation and a variety of practical steps that can be pursued that may help to size PETs' benefits relative to drawbacks. The evaluator can:

- Define data they consider to be sensitive.
- Identify all business processes and information systems that would be in scope for potential use of PETs; this will help define the boundary.
- Identify any instances of an information system or person with access to unneeded data elements.
- Evaluate different kinds of PETs to determine whether any could reduce the amount of sensitive information to be accessed, stored, or transmitted within the boundary while still allowing information systems and personnel to fulfill their duties effectively.
- Identify instances in which duplicate copies of records are maintained by in-scope systems or personnel for resiliency, testing, or other purposes.
- Work with the business, legal, operations, and technology teams to identify costs and other potential drawbacks of implementing PETs.
- Compare drawbacks against the qualitative factors and quantitative benefits calculated.

Conclusion

As payment systems and financial institutions evolve to adopt new payment methods and leverage new technologies, the threat landscape continues to evolve as well, presenting increased threats to data integrity and privacy. Consumers care deeply about the privacy of their financial information. Data breaches and similar incidents that expose sensitive personal information reduce trust in financial institutions, harm financial institutions' brand image, and generate additional societal costs. There is growing interest in exploring PETs to enable privacy-enhancing payment processing techniques while still allowing all parties in the payment process to fulfill their duties. The benefits of PETs can be compared against their drawbacks when evaluating their feasibility. These benefits can be qualitative or quantitative, and some practical steps can be taken to begin a benefits assessment. PETs may allow more payment activities to take place with fewer sensitive data elements accessible to people and systems.

The framework proposed here offers a method to quantitatively measure the related risk-reduction benefits.

References

(1) Federal Trade Commission (2002) "How To Comply with the Privacy of Consumer Financial Information Rule of the Gramm-Leach-Bliley Act," available at: https://www.ftc.gov/system/files/documents/plain-language/bus67-how-comply-privacy-consumer-financial-information-rule-gramm-leach-bliley-act.pdf (accessed May 1, 2024).

(2) General Data Protection Regulation (GDPR) (2018) "Definitions," available at: https://gdpr-info.eu/art-4-gdpr/ (accessed May 6, 2024).

(3) Fazlioglu, M. (2023) "Executive Summary - Privacy and Consumer Trust," available at: https://iapp.org/media/pdf/resource_center/privacy_and_consumer_trust_report_summary.pdf (accessed May 1, 2024).

(4) Anant, V., Donchak, L., Kaplan, J., and Soller, H. (2020) "The Consumer-Data Opportunity and Privacy Imperative," available at: https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/the-consumer-data-opportunity-and-the-privacy-imperative (accessed May 1, 2024).

(5) Petrosyan, A. (2024) "Number of Data Compromises and Impacted Individuals in U.S. 2005-2023," available at: https://www.statista.com/statistics/273550/data-breaches -recorded-in-the-united-states-by-number-of-breaches-and-records-exposed/ (accessed May 1, 2024).

(6) Donovan, S. (2017) "Preparing for and Responding to a Breach of Personally Identifiable Information, Executive Office of the President Office of Management and Budget Memorandum," available at: https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2017/m-17-12_0.pdf (accessed May 1, 2024).

(7) Bryan Cave Leighton Paisner Law (2014) "Data Breaches at a Glance," available at: https://www.bclplaw.com/a/web/2288/Data-Breaches-At-A-Glance.pdf (accessed May 1, 2024).

(8) Shah, K. (2023) "How to Regain Consumer Trust After a Data Breach," available at: https://business.yougov.com/content/46415-how-to-regain-consumer-trust-after-a-data-breach (accessed May 1, 2024).

(9) Caporal, J. (2024) "What Customers Want from Banks: Online Banking Trends and Consumer Priorities," available at: https://www.fool.com/the-ascent/research/digital-banking-trends/ (accessed May 1, 2024).

(10) Reimsbach-Kounatze, C., Reynolds, T., Girot, Clarisse (2023) "Emerging Privacy-Enhancing Technologies Current Regulatory and Policy Approaches," available at: https://doi.org/10.1787/bf121be4-en (accessed May 1, 2024).

(11) Bank of England (2023) “The Digital Pound: Technology Working Paper,” available at: https://www.bankofengland.co.uk/-/media/boe/files/paper/2023/the-digital-pound-technology-working-paper (accessed May 1, 2024).

(12) Digital Dollar Project (2023) “Revisiting the Digital Dollar Project’s Exploration of a U.S. Central Bank Digital Currency: White Paper 2.0,” available at: https://digitaldollarproject.org/wp-content/uploads/2023/01/DDP-Whitepaper-2.0_2023.pdf (accessed May 1, 2024).

(13) Mascelli, Jillian (2023) “Data Privacy for Digital Asset Systems,” Finance and Economics Discussion Series 2023-059. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2023.059 (accessed May 14, 2024).

(14) Asrow, K. and Samonas, S. (2021) “Privacy Enhancing Technologies: Categories, Use Cases, and Considerations,” available at: https://www.frbsf.org/wp-content/uploads/Privacy-Enhancing-Technologies_FINAL_V2_TOC-Update.pdf (accessed May 1, 2024).

(15) Blake, M., McWaters, J. and Galaski, R. (2019) “The Next Generation of Data- Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value,” available at: https://www3.weforum.org/docs/WEF_Next_Gen_Data_Sharing_Financial_Services.pdf (accessed May 13, 2024).

(16) Mastercard (2024) “Privacy Enhancing Technologies,” available at: https://b2b.mastercard.com/media/z0pnu32l/privacy-enhancing-technologies-white-paper-final.pdf (accessed May 3, 2024).

(17) Bischoff, P. (2023) “Financial data breaches accounted for 232 million leaked records across 2,260 data breaches - Comparitech,” available at: https://www.comparitech.com/blog/vpn-privacy/financial-data-breaches/ (accessed May 6, 2024).

(18) Board of Governors of the Federal Reserve (2022) “The Federal Reserve Payments Study: 2022 Triennial Initial Data Release,” available at: https://www.federalreserve.gov/paymentsystems/fr-payments-study.htm (accessed May 3, 2024).

(19) IBM Security (2023) “Cost of a Data Breach Report,” available at: https://www.ibm.com/resources/downloads/cost-data-breach-report (accessed May 1, 2024).

(20) Bank of England (2024) “Response to the Digital Pound: Technology Working Paper,” available at: https://www.bankofengland.co.uk/paper/2024/response-to-the-digital-pound-technology-working-paper (accessed May 1, 2024).

(21) Mixson, E. (2021) “The Pros and Cons of Privacy-Enhancing Technologies (PETs),” available at: https://www.cshub.com/executive-decisions/articles/the-pros-and-cons-of-privacy-enhancing-technologies-pets (accessed May 6, 2024).