Canadian Journal of Law and Technology

Volume 17 | Number 2 Article 5

12-1-2019

Privacy Law Issues in Public Blockchains: An Analysis of Blockchain, PIPEDA, The GDPR, and Proposals for Compliance

Noah Walters
Faculty of Law, Osgoode Hall Law School

Privacy Law Issues in Public Blockchains: An Analysis of Blockchain, PIPEDA, the GDPR, and Proposals for Compliance

Noah Walters*

TABLE OF CONTENTS

---

* JD / MBA, University of Ottawa Faculty of Law. I would like to thank Professor Elizabeth Judge at the University of Ottawa's Faculty of Law for her sincerity, consideration, and insight that guided me through the development of this research. Her stimulating questions and commentary consistently challenged me to improve.

Conclusion

Proponents of blockchain proclaim that the technology's greatest innovation is trust.[1] Blockchains create trust by serving as an indispensable ledger (a central point of truth), for all stakeholders to a transaction. Instead of companies managing and reconciling records of the same transaction in privately held databases, both sides of a transaction are recorded simultaneously on a shared ledger — the blockchain. As a result, the crypto economic environment is characterized by the decentralized coordination of business processes and transactions. Proponents of crypto-economics regard decentralized coordination as an opportunity for new forms of economic innovation, forms designed to increase value for individuals and decrease the power of intermediaries.[2] Blockchains enable decentralized coordination by increasing the transparency of information between parties. Transparency allows the blockchain network to police itself by allowing users to collectively verify the legitimacy of every network transaction.[3] Agents to a transaction can presume fair play in this system because the transparency, security, and immutability of data should theoretically lead to increased accountability for all participants.[4]

But while transparency can facilitate decentralized business relationships by allowing individuals and businesses to trust in the network itself (rather than the intermediary), unchecked transparency can pose a paradoxical threat to the

---

1 Jason Leibowitz, "Blockchain's Big Innovation is Trust, Not Money" *CoinDesk* (23 May 2016), online: < www.coindesk.com/blockchain-innovation-trust-money/ > ; "The Trust Machine" *The Economist* (31 October 2015), online: < www.economist.com/leaders/2015/10/31/the-trust-machine > ; McKinsey & Company, "How blockchains could change the world" (May 2016), online: < www.mckinsey.com/industries/high-tech/our-insights/how-blockchains-could-change-the-world > .

2 Christian Catalini & Joshua S. Gans, "Some Simple Economics of the Blockchain" (2016) Rotman School of Management Working Paper No 2874598, online: < papers.ssrn.com/sol3/papers.cfm?abstract_id = 2874598 > .

3 Primavera Di Filippi, "The Interplay Between Decentralization and Privacy: The Case of Blockchain Technologies" (2016) 7 J. Peer Production: Alternative Internets 5 at 5, online: < papers.ssrn.com/sol3/papers.cfm?abstract_id = 2852689 > [Di Filippi].

4 Immutability is not absolute; transactions can be reversed by blockchain forks. For more on this technical circumstance see Amy Castor, "A Short Guide to Bitcoin Forks" *CoinDesk* (16 May 2017), online: < www.coindesk.com/short-guide-bitcoin-forks-explained/ > .

privacy rights of network participants. The purpose of this research is to identify how blockchain technologies may clash with, and reconcile, privacy legislation. The research identifies the legislative issues related to the data subjects' privacy rights, and proposes technical and policy oriented solutions for privacy law compliance. The research is organized into four parts. Part one defines blockchain technology and distinguishes between blockchain forms to identify the nuances of public, private, and hybrid blockchains. Part two outlines the primary privacy law challenges to blockchain as a database and a medium of exchange, using Canada's PIPEDA legislation and the European Union's GDPR as a benchmark for legal analysis. Part three offers a non-technical primer of privacy-centric technologies designed to facilitate the compliant processing of data on public blockchains; this section discusses analytical techniques used to identify users on public blockchains. Part three also explores the implications of privacy-centric technologies as a solution to legislative compliance. Part four offers policy recommendations designed to complement proposed technical safeguards and generate a system of accountability in a network characterized by anonymity.

BLOCKCHAIN TECHNOLOGY: DISTINGUISHING BETWEEN PUBLIC, PRIVATE, AND HYBRID CHAINS

Blockchain can be explained as a digital infrastructure that serves two major and associated functions. First, blockchain is a medium of exchange that enables peer to peer transactions without the need for an intermediary. Second, blockchain is a database that serves a book keeping function for recording and verifying every transaction made on the blockchain in real time. Blockchain actions begin with someone creating a transaction, which can involve cryptocurrencies, smart contracts[5], records, digital representations of real world assets,[6] or other sets of information issued from a pseudonymous address (known as the public key address).[7] When a transaction is issued, the transaction request is broadcasted to all the computers in the network (these computers are referred to as nodes)[8]. The 'miner' nodes work to validate the transaction by solving complex algorithms designed to verify a transaction's legitimacy before appending it to the blockchain.[9] The mining process begins

---

5 A smart contract is a computerized self-executing protocol that enforces the execution of a predefined contract in a real-time manner: Nick Szabo, "The Idea of Smart Contracts" (1994), online: < www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html > .

6 Oliver Dale, "What is Tokenization? Real-World Assets on the Blockchain" *Blockonomi* (31 July 2018), online: < blockonomi.com/tokenization-blockchain/ > .

7 Stan Sater, "Blockchain and the European Union's General Data Protection Regulation: A Chance to Harmonize International Data Flows" (2017) at 19, online: < papers.ssrn.com/sol3/papers.cfm?abstract_id = 3080987 > [Sater].

8 Satoshi Nakomoto, "Bitcoin: A Peer-To-Peer Electronic Cash System (2008) at 1, online: < bitcoin.org/bitcoin.pdf > .

with miners bundling pending transactions to create a new block.[10] Each block has a header that contains the timestamp of the block, which includes a cryptographic hash of the block's items.[11] Along with the new block's cryptographic hash, the header also includes a reference to the previous block's hash, which creates a chain of verified transactions (the 'block' 'chain').[12] Every node on the network has a full replication of the valid transactions on the blockchain that dates back to the first block (the genesis block).[13] Once a block is successfully added to the chain, anyone on the network can query the transactions. Information is considered reliable because data recorded on each block is agreed upon by a majority of the network participants. Details of the data to a transaction and the rights of network participants vary by blockchain type and use case.

There are three types of blockchains: (1) public blockchains, (2) private blockchains, and (3) hybrid blockchains. Each type of blockchain is defined by its consensus mechanism, which is characterized by the way nodes participate to verify and record transactions.

Public blockchains

Public blockchains are considered "permissionless" in that data stored on the blockchain is open source and accessible by anyone (you *do not* need permission to participate). Here anyone can operate a node on the blockchain network, and every node has the capacity to read and write a transaction. Blockchain records are verified when 51% of nodes reach consensus. The identities of users and nodes on a public blockchain are pseudonymous; however, certain elements of transaction data are inherently transparent. As previously mentioned, a degree of data transparency is required so that network participants can trust the information stored on the blockchain. Primavera de Filippi explains that there are two layers of data in public blockchains: content data and protocol data.[14] Content data includes the terms and information of a specific transaction, whereas the protocol data includes contextual information about the transaction (i.e. the metadata).[15] Public blockchains are capable of privacy at the content

---

9 *Ibid.* at 3.

10 *Ibid.*

11 *Ibid.* A hash function takes an input (or 'message') and returns a fixed-size alphanumeric string that serves as the digital fingerprint of that input. It is a pseudonymization technique in cryptography used in storing and sharing information for efficiency.

12 *Ibid.*

13 Shaun S. Amuial, Josias N. Dewey & Jeffrey R. Seul, *The Blockchain: A Guide for Legal and Business Professionals*, eds. (Eagan, MN: Thomson Reuters, 2016).

14 Di Filippi, *supra* note 3 at 14.

15 Metadata (or "data about data") is everything about a piece of information, apart from the information itself. The content of a message is not metadata, but who sent it, when, where from, and to whom, are all examples of metadata. Legal systems often protect content more than metadata: for instance, in the United States, law enforcement needs a

level by way of encryption.[16] However, protocol data must be transparent for the blockchain to function.

Private blockchain

Private blockchains are considered "permissioned" because those who operate nodes on a private blockchain must be granted access to the blockchain by an administrator. There is only one administrator who has the capacity to read and write transactions to a private blockchain, and that is the owner of the blockchain. Thus, private blockchains are considered the most centralized version of blockchain technology because they are typically used as a tool by individual firms for internal processes. The administrator of a private blockchain protects information by maintaining control over the rights of network participants. Thus, degrees of transparency and data types are under the control of one data processor. All participants in a private blockchain are identifiable because the administrator must grant participants network access

Hybrid blockchains

Hybrid blockchains (or consortium chains) are controlled by a select group of administrators who have agreed to set terms that govern consensus.[17] They are public in that independently owned nodes must reach consensus for data to be validated, and they are private in that only those granted access to the blockchain can perform transactions. Only select users have the capacity to write transactions; however, reading rights can be designed as either public or private.[18] Like private chains, member nodes of consortium chains are also easily identifiable.

PRIVACY LAW CHALLENGES TO CRYPTO ECONOMIC DEVELOPMENT: THE CASE OF PUBLIC BLOCKCHAINS

A public blockchain is arguably the most disruptive form of the technology because of the highly decentralized nature of economic relations that the network enables. In contrast, private and hybrid blockchains embrace a degree of centralization common to traditional economic systems — this is demonstrated by the powers of control that groups or single administrators have over the rights of participants in the network (i.e. a user's ability to read information or execute

---

warrant to listen to a person's telephone calls, but claims the right to obtain the list of who you have called far more easily. However, metadata can often reveal a great deal, and will often need to be protected as carefully as the data it describes. See especially Electronic Frontier Foundation, "Metadata" (accessed 19 July 2019), online: < ssd.eff.org/en/glossary/metadata > .

16 Di Filippi, *supra* note 3 at 1.

17 Vitalik Buterin, "On Public and Private Blockchains" *Ethereum Blog* (2015), online: < blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/ > .

18 *Ibid.*

transactions). It is likely that the quasi-centralized character of private and hybrid blockchains generate discernable accountabilities in business relations that render such networks more easily reconcilable with existing privacy legislation. The following legislative review focuses on the implications of public blockchains because of the technology's distinctively decentralized functionality. This section is organized by jurisdiction, focusing first on privacy legislation in Canada, and second on privacy legislation in the European Union.

Canada

The Personal Information Protection and Electronic Documents Act (PIPEDA), sets the ground rules for how private sector entities handle personal information in Canada. Personal information is defined by PIPEDA as "information that on its own or combined with other pieces of data, can identify *you* as an individual."[19] Businesses who conduct operations atop blockchain infrastructure will need to comply with PIPEDA because the metadata necessarily ingrained in public blockchain transactions may constitute personal information. While context dependent, metadata will likely constitute personal information in the case of public blockchain transactions because it may reveal where transactions are sent from, who they are sent to (not necessarily the name of the recipient, but the address of the recipient), how much money was sent, and at what time. In *Gordon v. Canada* the Federal Court held that information is about an identifiable individual if it "permits" or "leads" to the possible identification of the individual, whether on the basis of that information alone, or when the information is combined with other information from other available sources.[20] This is corroborated by the Office of the Privacy Commissioner of Canada, which lists "financial information" as an example of personal information.[21] The same rules apply to every piece of software deployed on the blockchain (e.g. smart contracts), which are typically designed to execute business operations for companies building decentralized applications.[22] The operations of this code are made publicly available to every node in the blockchain network as "bytecode," which can be reverse engineered to reveal the same transactional information as the metadata in peer-to-peer transactions. To illustrate this point, the UK Government Chief Science Advisor explains that "chains of transactions may be traced throughout the Bitcoin blockchain to link,

---

19 > *Personal Information Protection and Electronic Documents Act*, R.S.C. 2000, c. 5, s. 2 at 4 [*PIPEDA*]. See also Office of the Privacy Commissioner of Canada, "Summary of Privacy Laws in Canada" (January 2018), online: < www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/02_05_d_15/ > [OPCC Summary].

20 Office of the Privacy Commissioner of Canada, *Metadata and Privacy: A Technical and Legal Overview* (Gatineau, Q.C.: Office of the Privacy Commissioner of Canada, October 2014) at 6, online: < www.priv.gc.ca/media/1786/md_201410_e.pdf > .

21 OPCC Summary, *supra* note 19.

22 Di Filippi, *supra* note 3 at 8.

for example, instances of bitcoin theft with specific attempts to withdraw bitcoins through exchanges."[23] While this analytical ability may serve a valuable purpose for fighting crime, it clearly demonstrates a privacy challenge in that identification of public blockchain users is possible.

Schedule 1 of PIPEDA outlines 10 guiding principles that a private enterprise must comply with to respect the privacy rights of individual consumers. The following section identifies each principle and expands on associated legal issues.

Principle 1: Accountability

An organization is responsible for personal information under its control and shall designate an individual or individuals who are accountable for the organization's compliance....[24]

PIPEDA's accountability principle was created to ensure organizational responsibility for protecting individual privacy rights. Designating accountability for public blockchain privacy infringement is a difficult task because public blockchains are not owned by a single entity. Rather, the ownership, and ultimately the ability to exercise control, over public blockchains becomes increasingly decentralized over time. To better understand the challenges associated with designating accountability, it is important to distinguish between the types of participants involved in a public blockchain network.[25] Borrowing from Dirk Zeutsch's ledger hierarchy, the relevant individuals and organizations in question generally fall under the following categories:

(1) The core group that develops the underlying software and principles that govern the public blockchain.[26]

(2) The owners of additional servers that run the public blockchain code for validation purposes (i.e. the nodes). [27]

(3) The companies that build decentralized applications on public blockchain infrastructure.

It is easy to presume that the core developers of the blockchain should be held accountable, because, after all, they compose the group that built the

---

[23] Mark Walport, *Distributed Ledger Technology: Beyond Blockchain* (London, U.K.: Government Office for Science, 2015) at 50, online: <assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf>.

[24] *PIPEDA*, *supra* note 19 at Sched.1, 4.10.

[25] As mentioned above, this section excludes the perspective of the public blockchains users who would constitute the 'data subjects' who disclose personal information when they make transactions on the blockchain.

[26] Dirk A. Zetzsche, Ross P. Buckley & Douglas W. Arner, "The Distributed Liability of Distributed Ledgers: Legal Risks of Blockchain" (2017) University of Hong Kong Faculty of Law Working Paper No 52 at 21, online: <papers.ssrn.com/sol3/papers.cfm?abstract_id=3018214>.

[27] *Ibid.*

system. Core developers set the rules that govern a public blockchain, and in many cases, they are the individuals who handle the maintenance of the network during its development.[28] However, as the community of owners on a public blockchain increases over time, the governance of the blockchain becomes increasingly decentralized (meaning that updates to the network must be approved by mutual consent from the community). Thus, public blockchain control is ultimately predicated upon the community of owners (that anyone can be a part of) who propose software updates and vote on suggested changes from other developers.[29] This makes it difficult to designate accountability to the core developers because while they approve changes, they can only do so with computational support from a majority of the community; their powers of control are not exclusive.[30]

The challenge of attributing accountability to core developers becomes more difficult when the core developers themselves are unidentifiable or unassociated with one another in the real world. The core developers behind the Bitcoin blockchain, for example, worked independently from one another, and the identity of the mastermind behind the network (Satoshi Nakomoto) remains unknown to this day. Nonetheless, in blockchain there is already a tendency towards centralization by powerful entities who have come to amass higher degrees of control than others. The primary example of this is Bitcoin, whereby miners are grouped in centralized communication pools, three of which control over 50% of the hash rate, while six pools control the 75% and the biggest individual pool controls 21.3%.[31] While the accumulation of hashing power will be unlikely to designate such mining groups accountable for privacy infringement, the existence of such groups challenges the purportedly

---

[28] On the Bitcoin blockchain for example, "decisions are made — or executed at least — by a team of core developers because only they have the technical permissions to accept submissions. Those core developers form, at least at first sight, Bitcoin's governance group in a narrower sense. Every adjustment to Bitcoin's governance structure must pass through the bottleneck of this small group of people." See Urs Gasser, Ryan Budish & Sarah Myers West, "Multistakeholder as Governance Groups: Observations from Case Studies" (2015) Berkman Center Research Publication No. 2015-1 at 8, online: < dash.harvard.edu/bitstream/handle/1/16140635/Berkman_2015-1-revision.pdf?sequence = 1 > .

[29] On Bitcoin, for any update "to become Active requires the mutual consent of the community. Those proposing changes should consider that ultimately consent may rest with the consensus of the Bitcoin users." See Bitcoin Core, "Bitcoin/bips" *GitHub* (accessed 19 July 2019), online: < github.com/bitcoin/bips/blob/master/README.mediawiki >

[30] On the bitcoin blockchain, for example, anyone who owns bitcoin also has owner abilities that allow such users to contribute to the decisions that determine the direction of the network. The more ownership of the network, the more influence you have over its direction.

[31] Roberta Filippone, "Blockchain and Individuals' Control Over Data in European Data Protection Law" (Thesis submitted in fulfillment of a LLM in Law & Technology, Tilburg University, August 2017), online: < arno.uvt.nl/show.cgi?fid = 143638 > at 33.

decentralized, democratic nature of community decision making that shields the core developers from accountability.

It is much less likely that owners of nodes will be designated accountable because anyone, anywhere, can download the entire history of data transactions that has taken place on a public blockchain. The fact that personal information may be gleaned from transaction records that node owners have downloaded poses an additional layer of privacy challenges related to principles of consent, limited collection, and limited use, disclosure and retention which will be discussed below.

The company that builds a decentralized application (DApp) on a public blockchain may be accountable in circumstances where a privacy breach can be traced to the accounts of the DApp team. However, the nature of 'decentralized' applications is designed to mimic the decentralized governance processes of public blockchains for a specific use case. In other words, the company's control over the network decreases over time as the applications community of users increases and becomes engaged in decision making. Therefore, the same issues for designating accountability with public blockchains exist with the decentralized applications built on top of them.

This is not to say that core developers will be absolved of liability from privacy-invasive design decisions. Rather, the nature of progressive decentralization will likely complicate fault attribution when design decisions are made by collective, unidentifiable parties, that span jurisdictions.

Principle 2: Identifying Purposes

The purposes for which personal information is collected shall be identified by the organization at or before the time the information is collected.[32]

PIPEDA's identifying purposes principle was created so that people could understand how their data was going to be used or disclosed. In the public blockchain context, while personal information may be derived from transactions, there is no explicit collection of information for an identifiable purpose. The only information included in the network is the metadata required to ensure network functionality. While personal information may be gleaned from this metadata, there is no mechanism in place to identify users when their data is being analyzed or for what purpose. There is also no disclosure of identifying purposes (or privacy risks) before the time information is collected. The concept of a public blockchain in general clashes with this principle because of the 'open-source' nature of the technology.

While public blockchain transparency creates inevitable privacy issues, some proponents of public blockchains argue that the technology enables the individual to exercise greater control over their personal information.[33] With a highly-decentralized architecture like a public blockchain, users would not have

---

[32] *PIPEDA*, *supra* note 19 at Sched. 1, 4.2.

[33] Angiva Banerjee & Karuna Pande Joshi, "Link Before You Share: Managing Privacy

to contend with the fear of data concentration and possible profiling by a third-party service because there is no central point of control.[34] However, the lack of centralized control does not reconcile the privacy issue that personal data can be obtained by any observer of the network. Therefore, one could also argue that, while decentralization reduces the threat of all data being controlled by a small group of central authorities (the Googles, Facebooks, and IBM's of the world), it does nothing to combat the threat of data manipulation by any singular unidentifiable entity who chooses to download a node.[35]

Principle 3: Consent

The knowledge and consent of the individual are required for the collection, use, or disclosure of personal information, except where inappropriate.[36]

Consent in the context of PIPEDA means informed and voluntary agreement with what is being done or proposed by the organization. It is only considered valid when it is reasonable to expect that someone can understand the nature, purpose and "*consequences* of the collection, use or disclosure to which they are consenting."[37] Consent can be either express or implied. Whether or not public blockchain users express or imply consent to the collection, use, or disclosure of personal information does not negate the fact that transactions executed over public blockchains become publicly available upon execution. Under PIPEDA, private enterprises are allowed to collect, use, and disclose personal information *without consent* when certain forms of personal information are made publicly available.[38] Examples of relevant forms listed under the regulation include: (a) personal information in a publicly available telephone directory, where the subscriber can refuse to have the personal information appear in the directory; (b) personal information that appears in a publicly available business directory, listing, or notice, where the collection, use, or disclosure relates to the purpose in

---

Policies through Blockchain" (2017) arXiv 1710.05363 at 2, online: <arxiv.org/pdf/1710.05363.pdf>.

34 Filippone, *supra* note 31 at 5.

35 While it is important to note that *PIPEDA*'s ten principles relate to individual data protection, there are similar privacy implications for businesses with respect to confidentiality. Public blockchain transparency in a competitive environment could challenge transactional confidentiality, which may dissuade firms from conducting business on public blockchains. This confidentiality issue could materialize in circumstances where firms have the capacity to track financial information from competitor transactions.

36 *PIPEDA*, *supra* note 19 at Sched. 1, 4.3.

37 Office of the Privacy Commissioner of Canada, *2017-2018 Annual Report to Parliament on the Personal Information Protection and Electronic Documents Act and the Privacy Act* (Gatineau, Q.C.: March 2018), online: <www.priv.gc.ca/en/opc-actions-and-decisions/ar_index/201718/ar_201718/#heading-0-0-3-1>.

38 Barbara McIsaac, Rick Shields & Kris Klein, *The Law of Privacy in Canada Vol 1* (Toronto: Thomson Reuters, 2016) (loose-leaf, release 2016), ch. 4 at 4-7.

which the information appears in the directory, listing, or notice; and, (c) personal information in a publication, including a magazine, book or newspaper, that is available to the public, where the individual has provided the information.[39] Much like in the case of directories and publications, public blockchain actors choose to send transactions on a platform that automatically publicizes the information of that transaction to all nodes in the network. Therefore, it is arguable that this functionality would trigger the 'publicly available' exception to PIPEDA's consent principle. If that is the case, then anyone with access to the blockchain would be able to acquire information from the transactions made by users of the blockchain without their consent. If the courts rule that the publication of transactions on the blockchain does not trigger the publicly available exception, then additional privacy law challenges arise, particularly challenges regarding section 4.3.8, which rules that "individuals may withdraw consent at any time."[40] Then, the blockchain feature of immutability directly conflicts with the individual's ability to withdraw consent.

One of the unique features of blockchains is that data stored on chain cannot be altered without acceptance of other nodes, which requires cooperation from more than half of the nodes for every transaction made.[41] When transactions are appended to the blockchain they are perpetually stored and build up over time. Even if a majority of nodes cooperate to implement a change, in order to remove an old transaction it would require nodes to verify the legitimacy of every affected transaction backwards, "un-build the entire blockchain block by block and then rebuild it afterwards."[42] This process requires extreme computational power and such an amendment would suspend all blockchain transactions until the reconstruction is complete. Thus, in practice the act of withdrawing consent, applied by the deletion or amendment of data on the blockchain, is very unlikely. This complication is exacerbated by the fact that there is no accountable entity to coordinate desired changes on public blockchains.

Principle 4: Limiting Collection

The collection of personal information shall be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means.[43]

The limiting collection was created to ensure that organizations only gathered relevant information about their data subjects. Applied to traditional economic systems the application of this principle is simple because

---

39 *Regulations Specifying Publicly Available Information*, SOR/2001-7, s 1.

40 *PIPEDA*, *supra* note 19 at Sched. 1, 4.3.8.

41 Matthias Berberich & Malgorzata Steiner, "Blockchain Technology and the GDPR: How to Reconcile Privacy and Distributed Ledgers" (2016) 2:3 European Data Protection L. Rev. 422 at 426 [Berberich].

42 *Ibid.*

43 *PIPEDA*, *supra* note 19 at Sched. 1, 4.4.

organizations have control over the data of their subjects. The peer-to-peer nature of public blockchain networks creates a greater level of complexity because individuals are the entities with control over data distribution. While only certain information is required to execute a public blockchain transaction, users have the ability to append additional information to the blockchain. One potential problem regarding limited collection persists in the circumstance where personal information about one individual is appended to the blockchain without their permission. Under such circumstances public blockchain features of transparency and immutability threaten the limiting collection principle in that potentially harmful actors could extract that information about users without their consent. There are currently no mechanisms in place to enforce this limited collection principle or safeguard the posting of illicit information.

Principle 5: Limiting Use, Disclosure, and Retention

Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by law. Personal information shall be retained only as long as necessary for the fulfilment of those purposes.[44]

As in the case of consent, the fact that histories of blockchain transactions can be downloaded by anyone creates an opportunity for harmful actors to dissect information and utilize it for purposes other than that for which it was originally provided. The purposes for which information may be provided can vary by use case and user intent. Some public blockchains were built to be replacements for fiat currency — as in the case of Bitcoin, and others were built to serve as a foundation for decentralized applications (DApps) — as in the case of Ethereum. In each case, data is stored on the blockchain differently. With Bitcoin's function being digital money, the data stored on the blockchain can be found in bitcoin transactions. Ethereum on the other hand is designed to act like a big computer that enables functionalities (in addition to the use of ether as a digital money). These functionalities are demonstrated by smart contracts. Smart contracts act like robot accounts by executing a piece of code when they receive a transaction.[45] The piece of code can be designed to automate a variety of different processes; it is through this functionality that DApps operate. The data collected by the DApp is stored in the smart contract, which is built on the Ethereum blockchain. All the data in every DApp on Ethereum is available to anyone who downloads an Ethereum node (referred to as the Ethereum Client).[46] Therefore, information can be gleaned from user accounts, simple cryptocurrency transactions (like a peer to peer transaction of bitcoin), and

---

[44] *Ibid.* at 4.5.

[45] Laurent Senta, "Where and How Application Data is Stored in Ethereum?" (accessed 19 July 2019), online: >www.singulargarden.com/blog/storage-and-dapps-on-ethereum-blockchain/<

[46] *Ibid.*

smart contracts. Neither Bitcoin nor Ethereum have enforcement mechanisms in place to protect user's personal information once it is made publicly available on the blockchain ledger. Generally, privacy solutions that have been proposed put the onus on the user to protect his/her information disclosed in transactions — examples of these solutions will be discussed in the final section of this research.

Principle 6: Accuracy

Personal information shall be as accurate, complete, and up-to-date as is necessary for the purposes for which it is to be used.[47]

On public blockchains there are native tokens (like Bitcoin for the Bitcoin blockchain, and Ether for Ethereum), and application tokens that give users access to the decentralized applications built atop public blockchain infrastructure. There is no mandated identification process for acquiring native tokens on public blockchains; accordingly, there is no way to ensure the accuracy of personal information. In the case of application tokens, companies looking to sell such tokens submit purchasers to requisite know your client (KYC) and anti-money laundering (AML) checks to comply with regulations in their jurisdiction. Such customer information is typically held off-chain (i.e. not coded into the public blockchain) in company-owned accounts. This information should be accurate to ensure compliance. However, transactional information is held on-chain, and once an application token circulates, anyone with a digital wallet can purchase or receive a token and gain access to the decentralized application. Therefore, while original KYC / AML records will likely reflect an accurate account of personal information for those who purchase the token from the issuing company, there is no mechanism in place to update information or ensure its accuracy over an extended period of time. This is because no exclusive controller exists to monitor all account related information and access points. A user who provides accurate information at the point of sale in 2016 may have committed a series of financial crimes in 2017, but would still be able to operate their public blockchain account under 2016 pretenses. Challenges to maintaining accurate accounts are exacerbated in circumstances where data is tampered with before a transaction is executed on the blockchain.[48] Under such circumstances, inaccurate information would be perpetually stored on the public blockchain ledger.

---

47 *PIPEDA*, *supra* note 19 at Sched. 1, 4.6.

48 Frank Hofman et al., "The Immutability Concept of Blockchains and the Benefits of Early Standardization" (November 2017) ITU Kaleidoscope: Challenges for a Data-Driven Society, online: < ieeexplore.ieee.org/document/8247004 > .

Principle 7: Safeguards

Personal information shall be protected by security safeguards appropriate to the sensitivity of the information.[49]

While transactional security is a major value proposition of public blockchains, the only layer of technical protection for personal information on most public blockchains is pseudonymous identity. Pseudonymous identity is different than anonymity. Anonymity means that it is impossible to link any given identifier to a specific identity; whereas a pseudonym merely refers to the use of an identifier to disguise the real identity.[50] Real identity can be uncovered by applying big data analytics to transaction histories stored on the blockchain ledger. The likelihood of discovery is dramatically increased when analytics include media-based correlating information.[51] Researchers have begun to develop new strategies for discerning real identities behind pseudonymous addresses, which would substantially weaken the privacy protection granted by pseudonyms in the digital world.[52] Various solutions to this problem have been proposed and will be discussed further in the final section of this research.

Principle 8: Openness

An organization shall make readily available to individuals specific information about its policies and practices relating to the management of personal information.[53]

The public blockchain ecosystem does not have any uniform policies or practices in place to reflect principles of openness. However, it is important that public blockchain operators adopt a system of best practices for managing personal information moving forward. Such practices should reflect disclosures of privacy risks at the point of token issuance (i.e. through ICO, airdrop, etc. . .). Sound practice should also identify where information is collected and shared; both on and off-chain.

But while token issuers may try their best to embrace open personal information policies and practices, the global and peer-to-peer nature of public blockchain operations poses a challenge to legislators. The global presence of token issuers, and open access to public blockchain networks, means that businesses can form anywhere, and market to customers across jurisdictions. Moreover, individuals have the power to purchase tokens from anyone,

---

[49] *PIPEDA*, *supra* note 19 at Sched. 1, 4.7.

[50] Di Filippi, *supra* note 3 at 11.

[51] Aidan Hyman, Interview with Chainsafe Systems, (August 7 2018 at 1:00pm) comment on linking pseudonymous accounts to user identities through public blockchains.

[52] One university simulated experiment used behavior based clustering techniques to reveal the identities of 40% of Bitcoin users: Elli Androulaki et al., "Evaluating User Privacy in Bitcoin" (Paper delivered at ETH Zurich, 2013) <eprint.iacr.org/2012/596.pdf>

[53] *PIPEDA*, *supra* note 19 at Sched. 1, 4.8.

anywhere, without much oversight from regulators. This makes it difficult to enforce regulations. While uniform principles of "openness" should be embraced across public blockchain networks, regulators currently have very little ability to ensure compliance with these practices.

Principle 9: Individual Access

Upon request, an individual shall be informed of the existence, use, and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.[54]

PIPEDA's individual access rights were originally intended to give individuals the ability to request information in corporate paper records or databases that wouldn't otherwise be accessible because companies owned the records and stored them privately. Once an individual requested access to records of their personal information, PIPEDA would allow the individual to make the case for "correcting" information if that information had changed since their previous recording. On public blockchains, the issue is not that individuals lack access to information, but that correction rights are difficult to establish because users cannot edit original records once they have been added to the blockchain ledger. In other words, while all records are accessible to all users for viewing, no individual can change one particular record. Moreover, public blockchains do not have a mechanism in place for users to request their own personal information, or to identify the context in which that information has been recorded. This poses additional challenges to accessibility when original records reflect inaccurate information. Such information could deceive observers without a wider context.

Principle 10: Challenging Compliance

An individual shall be able to address a challenge concerning compliance with the above principles to the designated individual or individuals accountable for the organization's compliance.[55]

Challenging compliance on public blockchains is difficult because the functionality of public blockchains is often at odds with data protection legislation. Features of transparency, decentralization, and immutability pose unresolved problems for PIPEDA's principles of accountability, information use, disclosure, collection, retention, and consent. However, public blockchain networks are open source by design, and regard information transparency as advantageous. The technology marks a clear departure from traditional social and economic norms by embracing transparency as a means of garnering trust. That said, the threat of bad actors exists, and the transparency inherent to public

---

[54] *Ibid.* at 4.9.

[55] *Ibid.* at 4.10.

blockchain networks creates substantial privacy risks emphasized by the legislative analysis above. Right now the only mechanisms in place for challenging compliance with the above fair information principles are through the privacy commissions that are not designed to account for the nature of public blockchain operations. Public blockchains do not have designated individuals for compliance.

European Union

The General Data Protection Regulation (GDPR) sets guidelines for the collection and processing of personal information of individuals within the European Union (EU).[56] The regulation applies to all organizations that process or control the personal data of members of the Union regardless of whether the processing takes place in the Union or not.[57] The GDPR differs from PIPEDA in that it takes a more stringent approach to protecting individual privacy rights by equipping data subjects with a greater degree of control over their own data through provisions such as the right to access[58] and right to be forgotten.[59] Amendments to this regulation also mandate that technical and organizational precautions be taken (such as privacy by design and appointing a data protection officer) to increase accountability for data processors and controllers.[60] The regulation's underlying philosophy seeks to increase individuals control over

---

[56] See European Commission, "Principles of the GDPR" (accessed 19 July 2019), online: .

[57] *General Data Protection Regulation*, Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, [2016] OJ L 119 at art. 3 [*GDPR*].

[58] Both PIPEDA and the GDPR grant individuals the right to access the personal information that the organization has about them; however, the GDPR introduces an additional right of "data portability" under Article 20. Data portability grants individuals the right to receive their personal data in a structured, commonly used and machine-readable format and to allow the individual to send that data to another data controller. See *GDPR*, *supra* note 57 at art. 20.

[59] *GDPR*, *supra* note 57 at art. 17 maintains an explicit right of erasure, whereas PIPEDA includes a 'qualified' right of erasure under principle 4.5. The principle in *PIPEDA*, *supra* note 19 at Sched. 1, 4.5 explains that "personal information shall be retained only as long as necessary for the fulfilment of those purposes." Article 17 includes an individual's right to require organizations to erase their data in a number of circumstances, particularly when there is a withdrawal of consent. In the GDPR this right extends to require the primary controller to take reasonable steps to inform other data controllers who have received the information of the withdrawal of consent.

[60] *GDPR*, *supra* note 57 at art. 4; 'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law; 'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

their data, while creating a more privacy-friendly environment for the processing of personal information.[61]

Similar to the case of PIPEDA, the GDPR defines personal information as that which contributes to the identifiability of a natural person.[62] An identifiable natural person is someone "who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."[63] Recital 26 states that "data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person."[64] On public blockchains, pseudonyms are the primary safeguard to mask user identity, and various studies have shown that big data analysis that combines blockchain data with external sources has the capacity to identify blockchain users.[65] Accordingly, it is likely that public blockchains and the organizations that utilize the technology, will need to comply with the GDPR. Moreover, while the GDPR does not explicitly comment on what degree of identifiability would result in data being deemed personal information, the only type of data precluded from GDPR regulation is anonymous data; therefore, it could also be implied that non-anonymous data, including pseudonymous data, is subject to the regulation.[66]

Presuming that public blockchains will fall under the purview of the GDPR, the following analysis identifies five features of the GDPR that serve as primary challenges to compliance:

Accountability

Article 5(2) of GDPR designates primary accountability to the *controllers* of personal data, and other direct obligations to the *processors* of personal data.[67] The GDPR defines a controller as "the natural or legal person, public authority,

---

[61] While legislative difference exist between PIPEDA and the GDPR, in recent years the Canadian government has embraced GDPR-like trends in their interpretation and appreciation of privacy laws. One House of Commons report in particular outlines a series of recommendations that embrace GDPR-like policies related to consent, data portability, a right to erasure, and privacy by design: see Canada, Parliament, House of Commons, Standing Committee on Access to Information, Privacy and Ethics, 42nd Parl., 1st Sess., *Towards Privacy by Design: Review of the Personal Information and Electronic Documents Act* (February 2018), online: < publications.gc.ca/collections/collection_2018/parl/xc73-1/XC73-1-1-421-12-eng.pdf > .

[62] *GDPR*, *supra* note 57 at art. 4.

[63] *Ibid.*

[64] *GDPR*, *supra* note 57 at recital 26.

[65] More on this in the analysis of privacy-centric technologies below.

[66] It is only the anonymized data to which the GDPR does not apply: *GDPR*, *supra* note 57 at recital 26.

[67] *GDPR*, *supra* note 57 at art. 5.

agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data,"[68] and a processor as "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller."[69] To illustrate this relationship, imagine Company A selling T-shirts through a website they've built using GoDaddy, while also using GoDaddy analytics to track customer engagement activity on their site. Here the controller is Company A, and the processor is GoDaddy. Data controllers are customers of data processors. The data controller is principally responsible for complying with data subject privacy requests, and data processors are obligated to only utilize data in ways that the controller instructs them to.[70]

The analogy to be drawn with public blockchains may exist in the distinction between providers of the infrastructure (blockchain protocols), and the providers of the decentralized applications built atop the infrastructure. However, public blockchain systems are designed so that full transparency of network metadata is accessible by anyone, anywhere. Accordingly, all nodes in the network technically process data because their purpose is to verify every transaction made. The big question for GDPR regulators here is what degree of control qualifies a node or developer as a data controller — the entity that is primarily accountable. Berberich and Steiner posit that the qualification of a data controller rests on whether the data holder has actual control over the personal information of users; then, either no node qualifies as a controller, or all nodes qualify.[71] Primavera Di Fillipi suggests that the regulation may be able to recognize the data subject as their own controller, and that "the responsibility of keeping data private merely shifts from the operator to the individual user."[72] Being that a fundamental characteristic of public blockchains is the lack of an identifiable entity in a position of control, regulators will likely face challenges designating an accountable entity in circumstances of privacy infringement.[73]

Data Minimization

Article 5(1)(c) of the GDPR explains that personal data shall be: "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (data minimization)."[74] As previously mentioned, decentralized coordination requires all nodes in a public blockchain network

---

68 *Ibid.* at art. 4.

69 *Ibid.*

70 MailControl, "Data Controllers and Processors" (accessed 19 July 2019), online: <www.gdpreu.org/the-regulation/key-concepts/data-controllers-and-processors/>.

71 Berberich, *supra* note 41 at 425.

72 Di Fillippi, *supra* note 3 at 15.

73 Cagla Salmensuu, "The General Data Protection Regulation and Blockchains" (2018) Liikejuridiikka at 12, online: <papers.ssrn.com/sol3/papers.cfm?abstract_id=3143992>.

74 *GDPR*, *supra* note 57 at art. 5.

to verify a transaction. The process requires nodes to access all previous transactions — essentially tracing the financial history of the parties to the transaction — to determine legitimacy.[75] The openness of this process can be opposed to the principle of data minimization because data from former transactions must remain accessible irrespective of whether the purpose of former transactions is still relevant. If data stored becomes an unnecessary byproduct of the public blockchains intended purpose, the blockchain may be in contravention of the data minimization principle. However, public blockchain contravention to the principle of data minimization depends on what the "purpose" of the information stored is. If the purpose of storing information is to ensure the security and accuracy of the whole, then the perpetual storage system of public blockchains may not violate the data minimization principle. For example, if a public blockchain system is applied to a government land registry, a comprehensive database of information that ensures the provenance of land transfers and ownership rights may not contravene the data minimization principle. Ultimately, compliance with data minimization will depend on the purpose for which data is processed on the public blockchain.

The principle of storage limitation

Article 5(1)(e) explains that personal data must be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed."[76] Accordingly, it is difficult to reconcile the public blockchain feature of perpetual storage with storage limitation. While there are exceptions to the storage limitation principle, none of them account for the current state of decentralized business operations that persist on public blockchains like Bitcoin or Ethereum.[77]

The right to be forgotten

Article 17 of the GDPR gives the data subject the authority to force the data controller to erase all data held by the controller about the data subject.[78] The right to be forgotten is exercised when the data subject withdraws his consent, when the data is unlawfully processed, or when the data subject objects to processing.[79] Each of these precursors to the right to be forgotten may pose issues for public blockchain operations. The withdrawal of consent first suggests the provision of consent. If consent is provided on public blockchains, it is likely implied because only the user has control over the processing of their own information. However, one must also consider the permanence of blocks to

---

[75] Filippone, *supra* note 31 at 30.

[76] *GDPR*, *supra* note 57 at art. 5.

[77] Personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes: *GDPR*, *supra* note 57 at art. 5(1)(e).

[78] *GDPR*, *supra* note 57 at art. 17.

[79] Sater, *supra* note 7 at 8.

evaluate the qualification of consent. Should both "permanence" and individual controls over the processing of their own information constitute lawful consent as previously understood for data protection purposes? Traditionally the answer to that question would be no; the GDPR is designed to empower individuals by giving them more control over their data. The permanence of data disables individual powers over their data over prolonged periods of time. Data unlawfully processed will also be stored on the immutable blockchain ledger.[80] Participants will be able to read the information deemed to be unlawful despite any subsequent attempts to correct or conceal such information. Lastly, a subject's objection to processing inevitably entails all data stored in the ledger because the blockchain's functionality, accuracy, and security relies on the immutability of transactional history. Of course, this begs the question of whether public blockchains leave any room for individuals to have control and consent over their data at all, after transactions have been made.

While this principle of the GDPR bears stark resistance to public blockchain operations, one possible exemption exists when there are "overriding legitimate grounds for processing."[81] Here the GDPR weighs the data subject's objection against the interests of a controller or third party. Although the spirit of the GDPR is such that individual data subject rights prevail over their controllers, the nature of public blockchains as a peer-to-peer system that operate without controlling entities could support an alternate theory. In the public blockchain context, one could argue that the legitimate interests of all independent users — to ensure network functionality and the sanctity of information, should override a single data subject's privacy request. Of course, this argument can only stand if the design of the public blockchain is such that certain privacy standards are met. However, if technological safeguards that generate privacy law compliance can be agreed upon, public blockchains could be perceived to substantiate the GDPR's prioritization of the data subject's rights because the nature of peer-to-peer operations inherent to public blockchains allow users to exercise more powers of control over their personal information. Without identifiable processors or controllers, individuals inevitably assume more responsibility over their own information on public blockchains. If privacy compliance can be reached with technical solutions or complimentary policy initiatives, the underlying public blockchain network could empower the data subject. A

---

[80] Processing is only lawful when: (1) the subject gives consent, (2) processing is necessary for the performance of a contract, (3) processing is necessary for compliance with a legal obligation, (4) rocessing is necessary in order to protect the vital interests of the data subject or of another natural person, (5) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, (6) processing is necessary for purposes of legitimate interests pursued by the controller or a third party: see Jacqueline van Essen & Vincent Wellens, "GDPR Series Part 6: Legal Grounds for Processing" *NautaDutilh* (blog) (16 February 2017), online: < www.e-nautadutilh.com/56/2608/landing-pages/news-item.asp?sid = f734514a-191f-49b0-8c60-2c1b34245d96 > .

[81] *GDPR*, *supra* note 57 at art. 17.

conversation on privacy-centric technology solutions and complimentary policy initiatives is included in the latter sections of this research.

Data Protection by Design and by Default (privacy by design)

Article 25 of the GDPR codifies the concept of privacy by design, by calling on data controllers to "implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed."[82] In other words, companies must develop their system architecture so that it protects user privacy rights from the onset. Public blockchains technically utilize privacy by design techniques like pseudonymization and encryption[83]; however, the implications of blockchain's decentralized control and distributed storage also conflict with privacy by design because blockchain systems are transparent, and do not minimize data or designate accountability.

PROPOSALS FOR SAFEGUARDING PRIVACY RIGHTS ON PUBLIC BLOCKCHAINS: PRIVACY CENTRIC TECHNOLOGIES

As is the nature of peer-to-peer systems, public blockchains facilitate transactions where the user exercises control over the transmission of data. However, most public blockchains can function only when a threshold amount of information is included in the transaction itself. This requisite amount is then broadcasted across all nodes in the network. The foundational privacy issue associated with public blockchain use results from the potential analysis of this information. As discussed in the introductory sections to PIPEDA and GDPR above, it is possible to analyze the metadata of public blockchain transactions and derive personal information that identifies users. This fundamental feature of public blockchains makes it difficult to accommodate the rights of data subjects related to use, disclosure, collection and consent. The following section identifies modern cryptographic proposals for safeguarding the privacy rights of users by anonymizing their identities on the blockchain. The section provides a non-technical primer of these technologies.

zk-SNARK (zero-knowledge Succinct Non-interactive Arguments of Knowledge)

Created by Eli-Ben Sasson et al as a privacy solution to public blockchain transactions (Bitcoin in particular), zk-SNARK technology allows users to hide their identities, transaction amounts, and account balances.[84] The technology

---

82 *GDPR*, *supra* note 57 at art. 25.

83 *GDPR*, *supra* note 57 at recital 78.

84 Eli Ben-Sasson et al., "Zerocash: Decentralized Anonymous Payments from Bitcoin" (Proceedings of the IEEE Symposium on Security and Privacy, San Jose, 17 May 2014), online: < zerocash-project.org/media/pdf/zerocash-oakland2014.pdf > [Ben-Sasson et al.].

does this by constructing a payment scheme between two parties that allows each party to prove to the other that they have a specific set of information, without revealing what that information is.[85] The most prominent use case for this technology is ZeroCash, the public blockchain cryptocurrency launched by Ben Sasson et al. The basis for this technical privacy solution is zero knowledge proof cryptography.

ZeroCash has received praise for its privacy-fulfilling technology, and criticism for its centralized control,[86] inefficiency, and poor scalability relative to other cryptocurrencies.[87] While zk-SNARK technology offers a viable privacy solution to some public blockchain transparency issues, the degree of anonymity that zk-SNARK enables raises other important regulatory concerns in that it creates greater barriers to "accountability" and "oversight."[88]

Mixing Techniques

Mixing services are offered through third parties to public blockchain transactions and can help users mitigate risks of identification. The service allows users to entrust an amount of coins to a pool operated by a party that 'mixes' coins from unassociated sources, thereby confusing the trail of cryptocurrency transactions, before redistributing the coins to the respective users.[89] Here users can send coins to each other in a way that hides the link between their old and new coins.[90] This helps mitigate the risk of identification because most de-anonymization techniques use "linkage attacks" that seek out connections between transaction inputs and outputs to identify users.[91] For mixing to be effective, the mixer needs many users with many coins to mix.[92]

Primary limitations to mixing services are that they typically charge fees, create delays, and become single points of failure for transactions themselves.[93] The mixer can steal the coins, can shut down, or can be hacked during the mixing

---

85 *Ibid.*

86 Joseph Young, "Merits & Limitations of ZCash: Thoughts of Experts" *CNN* (29 October 2016), online: < www.ccn.com/merits-limitations-zcash-experts-thoughts/ > .

87 Yuncong Zhang et al., "Z-Channel: Scalable and Efficient Scheme in Zerocash," (2018) 86 Computers & Security 112.

88 Ben-Sasson et al., *supra* note84.

89 *Ibid.*

90 Steven Goldfeder et al. "When the Cookie Meets the Blockchain: Privacy Risks of Web Payments via Cryptocurrencies" (2017)arXiv 1708.04748, online: < arxiv.org/pdf/1708.04748.pdf > .

91 Danny Yang, Jack Gavigan & Zooco Wilcox-O'Hearn, "Survey of Confidentiality and Privacy Preserving Technologies for Blockchains" *R3* (2016) online: < z.cash/static/R3_Confidentiality_and_Privacy_Report.pdf > .

92 *Ibid.*

93 Tim Ruffing, Pedro Moreno-Sanchez and Aniket Kate, "CoinShuffle: Practical Decentralized Coin Mixing for Bitcoin" (Proceedings, Part II of the 19th European Symposium on Research in Computer Security, Wroclaw, Poland, September 7-11,

period.[94] There is additional risk if the mixer keeps logs of the transactions it mixes because they could be analyzed to reveal precise linkages between participants that could then be de-anonymized to reveal personal information. Moreover, the technical nature of mixing is not impregnable, and vulnerability will likely increase as analytical abilities develop. One study indicates that mixing methods (based on CoinJoin — the most common mixing technique do not always provide anonymity gains for users.[95]

Ring Confidential Transactions (Ring CTs)

Ring Confidential Transactions are built using ring signature technology, which was first introduced by Rivest, Shamir and Tauman in 2001.[96] The original function was to enable government officials to leak secret information without revealing who disseminated the information.[97] The technology was amended for public blockchain use by bitcoin core developer Gregory Maxwell and formally produced in the major cryptocurrency Monero in 2015.[98] Ring CTs are cryptographic digital signatures that protect sender privacy by obscuring the input side of the transaction amongst arbitrary users — making it computationally infeasible to determine who the signer of a transaction is.[99] In other words, a message signed with a ring signature is endorsed by someone in a random group of people, but the actual signer is not distinguishable among the group.[100] In Monero, such ring signatures are composed with outputs from the real sender's address, alongside a number of decoy addresses known as "mixins."[101] Monero provides an additional layer of privacy to this technology by incorporating 'stealth addresses' into the payment scheme.[102] A stealth

---

2014) online: <link.springer.com/book/10.1007/978-3-319-11212-1> and <petsymposium.org/2014/papers/Ruffing.pdf> [Ruffing et al.].

94 *Ibid.*

95 Felix Konstantin Maurer et al., "Anonymous CoinJoin Transactions with Arbitrary Values" (Paper delivered at IEEE International Conference on Trust, Security, and Privacy in Computing, Sydney, 1 August 2017), online: <ieeexplore.ieee.org/abstract/document/8029483>.

96 Bryan Curran, "What Are Ring Signatures? Providing Privacy for Cryptocurrency" (7 July 2018), online: <blockonomi.com/ring-signatures/#Background_of_Ring_Signatures> [Curran].

97 *Ibid.*

98 *Ibid.*

99 Henning Kopp et al., "Design of a Privacy-Preserving Decentralized File Storage with Financial Incentives" (Paper delivered at IEEE European Symposium on Security and Privacy Workshops, Paris, 26 April 2017), online: <ieeexplore.ieee.org/document/7966965> [Kopp et al.].

100 Curran, *supra* note 96.

101 Justin Ehrenhofer, "Response to 'An Empirical Analysis of Traceability in the Monero Blockchain', Version 2," (29 March 2018), online: <getmonero.org/2018/03/29/response-to-an-empirical-analysis-of-traceability.html>.

address requires the sender to create "random one-time addresses for every transaction on behalf of the recipient."[103] Stealth addresses hide the receiving addresses of transactions, and Ring CTs hide the sender's identity and the transaction amounts. Before a 2017 update to the Monero network, one study revealed that a "chain-reaction" analysis could be used to deduce the real sender's identities at 85% accuracy.[104] The update has since mitigated this de-anonymization technique by adding more mixins to Monero transactions. The rate of deductability (i.e. de-anonymization) is now close to 0%.[105]

While Monero provides valuable privacy controls, its degree of anonymity coupled with decentralization raises similar concerns to ZeroCash in that it enables criminal activity, and makes it extremely difficult to determine accountability and provide oversight. Moreover, the underlying privacy-centric technologies that Monero utilizes are only applicable (as of the date of this paper) to public blockchains as alternative payment mechanisms — rather than a blockchain designed for running decentralized applications like Ethereum.

Storing Personal Data Off-Chain

One simple approach to protecting privacy on public blockchains is to store all personally identifiable information off chain. This approach provides privacy by restricting access to the data, but trusted third parties (TTP) are required. Parties storing personal information off chain would create a hash of the transaction details on chain.[106] By hashing the data, observers would be unable to glean personal information from the transaction itself; however, this process requires both counterparties to verify that the hash of data on chain matches with their records off-chain.[107] This allows public blockchain users to keep the details of their transactions private from other blockchain participants, but it also undermines many of the advantages of using the blockchain as a shared ledger.[108] After all, what would be the point of using a shared ledger if network participants need to reference their own off-chain ledgers each time they execute a transaction.

---

[102] *Ibid.*

[103] Kopp et al., *supra* note 99.

[104] Malte Moser et al., "An Empirical Analysis of Traceability in the Monero Blockchain" (2017) arXiv 1704.04299, online: < arxiv.org/pdf/1704.04299/ > [Moser et al.].

[105] *Ibid.*

[106] A hash function takes an input (or 'message') and returns a fixed-size alphanumeric string that serves as the digital fingerprint of that input.

[107] Ruffing et al., *supra* note 93.

[108] *Ibid.*

Implications of the Privacy-Centric Technology Approach to PIPEDA and GDPR Compliance

Modern cryptographic proposals for safeguarding privacy rights may achieve compliance through anonymization. The anonymization approach presumes that public blockchain operations will comply with, or be exempt from, legislation like PIPEDA and the GDPR if user data cannot be linked to user identity. Here the logic follows that if data stored is completely anonymous, then such data cannot be considered "personal information" to be processed or gleaned from parties subject to the legislation. Recital 26 of the GDPR supports the anonymization approach to privacy when it states:

> Principles of data protection should (. . .) 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. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.[109]

The Office of the Privacy Commissioner in Canada takes the same approach to anonymous information, explaining that anonymous information should not be considered personal information, "as long as it is not possible to link that data back to an identifiable person."[110] Technically speaking, the anonymization of data should exempt public blockchain operations from these regulations. However, the reality of technological innovation is such that the privacy-centric technologies that enable anonymity become outdated and incapable of masking user identity over prolonged periods of time. History has proven this to be true with Bitcoin, which many originally believed to be anonymous, and more recently with Monero, where one study was able to uncover up to 85% of sender identities (before the 2017 update).[111] Moreover, the standalone technical solutions that embrace anonymity fail to account for other important regulatory considerations that arise from a transactional paradigm that is decentralized. As previously mentioned in the analysis of PIPEDA and the GDPR, the decentralized, 'peer-to-peer' nature of operations on public blockchains makes it extremely difficult to designate accountability for instances of privacy infringement. For example, while the inability to identify users and read transactional information should reconcile issues of collection, use, storage minimization, and the right to be forgotten (because there would be no *personal* information to collect, use, store, or forget), it may also burden regulators seeking to protect society from financial crimes like money laundering, terrorist financing, and tax evasion. Accordingly, a transactional paradigm that is both

---

[109] *GDPR*, *supra* note 57 at recital 26.

[110] OPCC Summary, *supra* note 19

[111] Moser et al., *supra* note 104.

anonymous and decentralized could generate a marketplace with little oversight and enforceability. To preserve the security of network participants and respect privacy rights, a practical approach to regulation will be one that enables accountability while maintaining transactional anonymity. Such an approach requires policy that complements the adoption of privacy-centric technology.

PROPOSALS FOR SAFEGUARDING PRIVACY RIGHTS ON PUBLIC BLOCKCHAINS: POLICY RECOMMENDATIONS

This section introduces four policy recommendations that seek to integrate accountability into a decentralized system that relies on transactional anonymity to protect user privacy. The purpose of these recommendations is to facilitate privacy-compliant public blockchain operations while at the same time offer a viable means for regulators to protect society from unlawful activities like financial crimes. The recommendations presume that the public blockchain ecosystem will require transactional anonymity and accountability to flourish; accordingly, privacy-centric technologies are an assumed tenet of the public blockchain ecosystem pursuant to the following recommendations.

Regulate Crypto Currency Exchanges

Participation in the current public blockchain ecosystem requires the use of cryptocurrencies. The only way to purchase cryptocurrencies initially is by using a cryptocurrency exchange. The exchange is a viable target for generating a means to hold people accountable because it serves as the primary entry and exit point to the public blockchain marketplace. A focused regulatory strategy that mandates exchanges to conduct operations that comply with privacy legislation could enable accountability in a system of anonymous transactions. Here exchanges would presume responsibilities similar to banks. Using an exchange would require participants to meet KYC and AML criteria, and the exchange would also keep records of user accounts and transactions (these records would be kept off-chain or on a private sidechain). Nodes to the public blockchain network would still exist, and third parties would still be able to download data, but privacy-centric technologies would shield users from bad actors seeking to take advantage of personal information. Exchanges would be the only third party with the ability to identify user accounts, and regulators would have to work with exchanges to police the network. While this may seem burdensome, the degree of digitization involved would likely mean that much of the regulatory activity would be automated; only flagged transactions would be brought to the attention of human regulators. One could look to the automated infringement notice systems in the copyright context as a benchmark for development. Such systems are used by universities to protect against illegal downloading and sharing of files. Another example derived from public-sector surveillance is the NSA's "PRISM" program. PRISM was the analytics program that flagged court approved terms when they appeared in data processed over the internet and in

telecommunications.[112] Despite the controversy surrounding the PRISM program in particular, regulators could implement similar analytical procedures tailored to cryptocurrency exchanges to flag suspicious activity. The program could scan the backend blockchain ledger and flag information based on suspicious transaction amounts, locations, and the criminal affiliations of users. Under this regulatory regime the exchange would only be held responsible for upholding regulatory requirements by design, and integrating the analytical program into their network so that regulators could track behind the scenes information that the program flags as noteworthy or suspicious. It is important to note however, that while a PRISM like algorithm may provide a degree of regulatory certainty, such backend access to information could also challenge the very purpose of public blockchains as a trusted means of decentralized coordination. For that reason, this recommendation could face significant resistance from the public blockchain community; a community born of a philosophy that prioritizes anonymity and the privacy of communications.[113]

Mandated Government Registration for Anonymous Accounts

This approach to regulation would require a government entity (or international body) to be the sole issuer of cryptocurrency wallets. A cryptocurrency wallet stores the public and private keys which public blockchain users must hold to receive or send cryptocurrencies. If regulators exercise control over the wallet distribution process, they could keep records of public blockchain user accounts and monitor activity on public blockchains accordingly. Users would still be capable of transacting peer-to-peer, and all transactional information would remain anonymous. Government control over wallet distribution would enable regulators to track wallet holder activity privately. Oversight of this process would have to remain behind closed doors. It would also likely require cooperation amongst nation states for information sharing.

Deem Public Blockchain Data Transfers "Consensual"

Both PIPEDA and the GDPR create an exemption clause for circumstances where the data subject consents to the processing of their personal information. Privacy regulators could approach public blockchains as systems where participation implies consent. Justification for this approach is twofold: (1) the database is inherently transparent and therefore participants recognize that transactions will be broadcasted to all nodes in the network, and (2) transactions occur peer-to-peer, meaning that the data subjects themselves exercise exclusive

---

[112] Zygmunt Bauman et al, "After Snowden: Rethinking the Impact of Surveillance" (2014) 8:2 Int'l. Political Sociology 121 at 123.

[113] For more on this philosophy see: Eric Hughes, "The Cypherpunk Manifesto" (9 March 1993), online: < www.activism.net/cypherpunk/manifesto.html > .

control over their transactions. In the case of PIPEDA, this argument may be more likely to succeed because the legislation also includes a clause deeming publicly available information accessible to third parties. Once a transaction is broadcasted across all nodes, it is made accessible (or publicly available) to any interested party. The case of the GDPR is more complicated because it does not reference the public domain as a factor in data processing. Rather, the GDPR explains that processing is allowed only when it is consented to by the data subject. Here a regulatory approach would rely on the release of a formal interpretation of GDPR consent for public blockchains, or a relaxed interpretation of consent with respect to public blockchain transactions. A relaxed interpretation could be substantiated by the fact that there may be no 'data processors' or 'data controllers' on public blockchains other than the users themselves. Users can be defined as both processors and controllers for all public blockchain transactions because there is no intermediary to facilitate the transactional procedure; there is only the network itself. To hold the network accountable could be considered unreasonable because it is simply a tool to facilitate peer to peer transactions.

If regulators deem public blockchain data transfers to be an inherently consensual process, regulation must instead focus on the analysis and use of data by third parties. Especially heavy penalties could be posited to deter malicious or invasive use of information gleaned from public blockchain operations to help counteract the vulnerabilities associated with a transparent system of transactions. Of course, there have long been problems with public records being used for nefarious purposes despite the penalties that already exist. Accordingly, it is uncertain whether heavier penalties will be effective. This approach would also require regulators to exempt public blockchains from the right to be forgotten, or at the very least amend the clause to accept 'corrected' information as a sufficient response. Information posted on a public blockchain that a data subject wants erased could be 'corrected' by indicating a desired edit to the information in a subsequent transaction. This approach to correction would not 'erase' the original transaction, rather it would serve an organizational function by providing more recent, accurate, information for observers.

Establish a 'Technically Sophisticated User' Clause

Introducing a technically sophisticated user clause would serve the function of easing privacy regulation principles for data processors and controllers. By easing privacy regulation in public blockchain circumstances, regulators would be able to remain passive, and only intervene in instances of intentional wrongdoing. Regulators could use a technically sophisticated user clause to justify their passive approach to privacy regulation in the public blockchain ecosystem. A passive regulatory approach can be justified by the fact that participation in this ecosystem requires a much higher degree of complexity and independence than the traditional transactional paradigm where intermediaries facilitate actions on our behalf. Creating a wallet and executing a cryptocurrency

transaction is a more complicated process than sending an e-transfer. Moreover, editing text into a transaction for purposes of information sharing requires a degree of technical sophistication that clearly supersedes that required for sending an email or writing a blog article. Just as accredited investors abide by different standards for participation in secondary markets, a technically sophisticated user status could justify a different approach to privacy regulation on public blockchains.

An important counter to this argument is that if the technology is rolled out with more mundane consumer transactions, individuals may not know how their personal information is being stored (in the same way that people typically do not internalize how accessible their data trails are in other contexts). Should peripheral technologies develop on blockchain technology that enable more simplified access to blockchain applications, the once 'sophisticated' early adopters will not be the only ones using the technology. This trend in technological adoption is commonplace; creating an email account is much easier today than it was in the 90s, as is accessing the internet. Accordingly, legislation that adopts a sophisticated user exemption is likely short-sighted. A more ubiquitous application of blockchain technology in the future may result in the technology being more invisible to those who use it.

CONCLUSION

Public blockchain features of decentralization and transparency conflict with contemporary privacy legislation like PIPEDA and the GDPR. Decentralization often renders the designation of accountability impossible, and transparency leaves the personal information of data subjects vulnerable. Both PIPEDA and the GDPR rely on accountable entities like processors and controllers to serve as both the champions of the regulation and the targets for infringement. This is complicated on public blockchains because the data subjects themselves often assume the roles of processors and controllers as transactional ability is conducted peer-to-peer. However, the primary privacy problem associated with the transparency and decentralization of the network is derived from the fact that the analytical ability exists to identify pseudonymous users and personal information associated with their transactions. While privacy-centric technologies like zk-SNARK's, Ring CTs, and mixing techniques, exist to generate anonymity for transactional information and accounts, this approach is not a standalone answer to our privacy needs. Anonymity without a complimentary regulatory response may exacerbate societal vulnerabilities by creating a marketplace with limited enforceability and oversight. While early blockchain adopters were looking for a system with less oversight and government intervention, a degree of regulation is likely required for the technology to achieve mainstream adoption. Privacy-centric technologies that provide anonymity require a complementary policy response to establish the foundation of a system that provides accountability for data processors and controllers. A grassroots approach to regulation borrows from the GDPR's

principle of privacy-by-design and encourages anonymous transactions, supported by regulatory oversight at the entrance and exit points of the public blockchain ecosystem. Anonymity will allow participants to utilize the network without concern for data theft or eavesdropping, whereas the backend ability to identify participants will deter bad actors from taking advantage of the system for unlawful purposes like financial crimes or illicit dark web activities. If regulators and developers work together, privacy law and privacy technologies can provide the answer to many of the legal issues and economic impediments facing mainstream public blockchain adoption.