2021

How Blockchain and AI Enable Personal Data Privacy and Support Cybersecurity

Stanton Heister
Portland State University

Kristi Yuthas
Portland State University, yuthask@pdx.edu

Citation Details

Heister, S., & Yuthas, K. (2021). How Blockchain and AI Enable Personal Data Privacy and Support Cybersecurity. In Blockchain Potential in AI. IntechOpen.

Chapter

How Blockchain and AI Enable Personal Data Privacy and Support Cybersecurity

Stanton Heister and Kristi Yuthas

Abstract

Recent increases in security breaches and digital surveillance highlight the need for improved privacy and security, particularly over users' personal data. Advances in cybersecurity and new legislation promise to improve data protection. Blockchain and distributed ledger technologies provide novel opportunities for protecting user data through decentralized identity and other privacy mechanisms. These systems can allow users greater sovereignty through tools that enable them to own and control their own data. Artificial intelligence provides further possibilities for enhancing system and user security, enriching data sets, and supporting improved analytical models.

Keywords: personally-identifiable data (PII), personal-data privacy, decentralized identity (DID), self-sovereign identity, cybersecurity, GDPR, zero-knowledge proofs

1. Introduction

The amount of personal data being collected is rapidly proliferating. Enterprises and governments use this data to profile individuals and to predict and control their attitudes and behavior. This can result in customized experiences, personalized services, and more efficient use of resources. It can also result in misinformation and exploitation by the entity that collected the data or by others that purchase or steal it. In response to increases in cybercrime and growing consumer concern, legislation to protect personal data is being proposed and implemented. Organizations trading in personal data face increasing costs associated with managing and securing data. They also face increasing risks that data will be misused or stolen, and that they will face legal or financial consequences, as well as damage to both their reputation and to relationships with customers and other stakeholders.

In this chapter, we explore how blockchain and artificial intelligence can offer solutions for protecting and securing personal data. Decentralized and federated identify systems provide users control over what, when and how much of their personal information can be shared and with whom. These systems can also reduce cybersecurity threats. Artificial intelligence complements blockchain-based privacy solutions by enabling users to better manage their data and by ensuring that data and models derived from the data are more accurate, fair, and reliable.

2. Personal data privacy

A foundational privacy issue facing information system developers and users is personal data privacy. Personally-identifiable data about clients, employees, prospects and other stakeholders may be regularly collected and stored in shared ledgers. Today, many organizations store private stakeholder data and even passwords in unencrypted form. Even when data are encrypted or anonymized, it may be possible to identify users unless well-developed cybersecurity processes are designed into data management systems. With frequent cybersecurity failures and increasing regulation, maintaining the privacy of personally identifiable information (PII) has become an issue of strategic concern for many organizations.

PII includes any data that can be traced back to a specific person, and can include individual items such as biometric data, social security numbers, phone numbers, or geolocation data. PII can also include data combinations, such as postal codes, birthdates, and gender, or behavioral data associated with one person. Organizations gather and store personal data about current and future customers and employees as well as about other stakeholders.

3. Cybersecurity and privacy breaches

Cybersecurity has become increasingly important for governments and businesses alike. Information security—one component of cybersecurity—focuses on protecting the integrity and privacy of data as it is captured, stored and used. The people, processes, and technology associated with data work in concert to create and maintain security.

Despite advances in security protocols and software, privacy breaches are on the rise. According to Risk Based Security's 2020 data breach report, "The total number of records compromised in 2020 exceeded 37 billion, a 141% increase compared to 2019" [1]. Personal records of system users are regularly compromised, and millions of these records, including names, emails and passwords, have been subject to data breaches, in many cases even including addresses, birth dates and financial information [1].

A data breach occurs from unauthorized access to an organization's database, enabling cyber hackers to steal sensitive personal information such as passwords, credit card numbers, social security numbers, and banking information [2]. These well documented breeches have had adverse consequences, including credit card fraud, and identity theft, which can have lasting negative effects on personal credit, often taking months, if not years, to remedy [2]. Some of the Largest, most recent cyber hacks include the 2013/14 breech of Yahoo's database by what is thought to have been a state-sponsored cyberattack, impacting over 3 billion users. The hackers collected consumers' names, email addresses, telephone numbers, dates of birth, hashed passwords and unencrypted answers to security questions.

In 2017, the credit reporting agency Equifax was subject to a cyberattack in which affected an estimated 143 million consumers. System administrators weren't aware of the suspicious activity for two months and did not report the breach for a full month after its discovery. It is believed that Equifax was breached by Chinese state-sponsored hackers engaged in espionage [3]. The collective financial impact to individual victims is not known, nor is it known what security and strategic damage was incurred by the state, but these cases highlight the potential risk when PII are housed in a centralized data base.

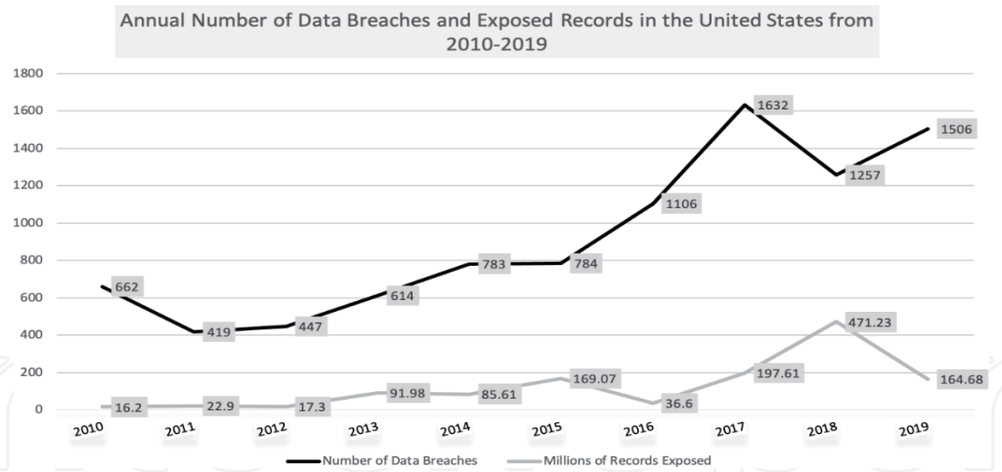

Figure 1.
Cybersecurity breaches and record exposure.

Most of the data gathered and stored are in the control of governments and corporations, which have gathered volumes of personal information that they are responsible for securing. At the same time, these organizations may be monetizing these datasets, either by using them to improve their own operations and offerings or by selling them to third parties. The volume of data generated and collected is increasing exponentially, enlarging the footprints of users. Data consolidators are able to link data elements across data sources and combine data in ways that were never anticipated by the parties that collected the information nor by the users that provided it.

Figure 1, which uses from data provided by Statista [4], shows the cost of amassing these large databases. Statista, a statistical research firm, tracks cybersecurity failures and trends. A recently published Statista report reveals that these events are increasing, especially in the past five years, underscoring the need to improve how data are secured. It should be noted that in 2020 a massive cyber breach by what is thought to be Russia could result in higher numbers for 2020 especially in the records exposed category as it is thought to be significant. The extent of the breach is still under investigation at the time of this publication.

4. Privacy regulations

The right to privacy is a considered to be basic human right in many parts of the world. That privacy may extend to individuals' right to control their own personal data. This right must be carefully defended as ownership and management of an individual's personal data can impact relationships with others and even the data-owner's identity [5].

Regulations governing how personal data are gathered and managed are rapidly being developed. The European Union has led the way in legislating privacy law through the General Data Protection Regulation (GDPR), passed in 2016. The law requires organizations, that gather personal data about EU citizens for transactions with EU member states, must carefully protect that data to ensure privacy.

In the US, the California Privacy Rights Act (CPRA) which expands on the 2018 California Consumer Privacy Act (CCPA), adopts many principles from the GDPR [6]. The CCPR is designed to provide residents of California the right:

1. to know what personal data is being collected

2. to know whether it is being sold or disclosed and to whom

3. to refuse the sale of their personal data

4. to access their personal data

5. to request that a business delete any personal data

6. not to be discriminated against for exercising their privacy rights [7]

At the federal level, the Consumer Online Privacy Rights Act (COPRA) was introduced in December 2019 by Democratic senators, led by Maria Cantwell. Although this bill has yet to pass, and previous federal privacy bills have failed, governmental bodies continue to pursue stricter laws for governing data [8].

Privacy laws directly affect how companies operate and will require firms that use consumer data to implement systems and operational practices that enable them to conform to these new regulations. Blockchain and Distributed Ledger Technology are uniquely positioned to help companies comply with existing and potential future regulation as it relates to personal property and data privacy.

5. Blockchain and privacy

Among the significant benefits of blockchain solutions is that they enable organizations to share data in ways not previously available, opening up possibilities for enhanced collaboration, improved operational efficiencies and expanded revenue. Questions about how to maintain privacy over the data are heightened in these environments because the data are stored in shared ledgers which may be accessible by multiple blockchain participants.

ConsenSys, a blockchain technology solutions company, in discussing the security of public blockchains, argues that "In reality, privacy is not a property of any blockchain. Rather, there are layers of privacy that can be applied to any blockchain..." [9]. Designers must carefully consider which parties are allowed to read and write transactions and how transactions are broadcast, validated, and stored. Additional issues relating to how permissions and security measures are updated and enforced are also important considerations. Decisions about who owns the data and how data can be used by organizations and computer applications further complicates privacy discussions [9].

5.1 Decentralized identity

Self-sovereign identity, a widely held view among blockchain proponents, holds that individuals should have control over their own identities and should have autonomy over how facets of identity are shared with others. Decentralized identity (DID) is a blockchain-enabled embodiment of self-sovereign identity that can profoundly improve the privacy and security of personal data.

DID refers to individual ownership of personal digital data relating to many elements of identity. Microsoft, which participates in defining DID standards, takes the perspective of the individual. "Currently, our identity and all our digital interactions are owned and controlled by other parties, some of whom we aren't even aware of [10]." Returning ownership of data to the individuals to whom the data pertains can provide benefits both to those individuals and to organizations that would otherwise be responsible for protecting the data.

Blockchain technology enables DID and provides a way for individuals to store their own data outside of the databases of the parties with whom they transact. Data are owned and controlled by these individuals and pointers to this data or metadata can be stored on the blockchain and can be used to verify the validity of claims the users make about their personal data. For example, a driver's license bureau might issue a driver's license to a user, which the user stores privately. When an insurance company or other party wishes to verify that the user is licensed, the user can present the license to a party such as an insurance company, and the party can independently verify the issuer and expiration date.

Anyone can create a DID. When this identity is first created, there is no information attached to it. Over time, the user could attach a driver's license or other identifying data to that DID. The process that a third party might use to verify that a particular person owns a DID, is similar to the process of validating that a person owns an email address. For example, an online gaming account can be attached to an email address. A party seeking to validate that a person was the owner of that account could send a private message, such as a security code, to the email address and ask the person to provide that code, something that only the person possessing the password for that email address could provide.

Unlike an email account, the DID would be owned and stored by a person rather than by an email service provider. The password, or private key, would also be secured by the owner. Personal information relating to the identity could be stored in an identity hub—an encrypted repository of personal data that is stored outside the blockchain, likely in a combination of phone, PC, and cloud data or offline storage devices [10]. Through the use of an identity hub, the person could control which pieces of information to share with an external party.

DIDs reduce the probability of unwanted correlation. The use of common identifiers—such as email addresses on different web sites—creates what is called a correlation problem. Correlation in this context means entities can, without a user's consent, associate information about a single identity across multiple systems. Email addresses utilize data on almost every website. When users provide the same email address on different sites—along with perhaps additional pieces of personal information like a phone number or physical address—they unknowingly enable a potential for correlation. In this case, entities can correlate that data across sites.

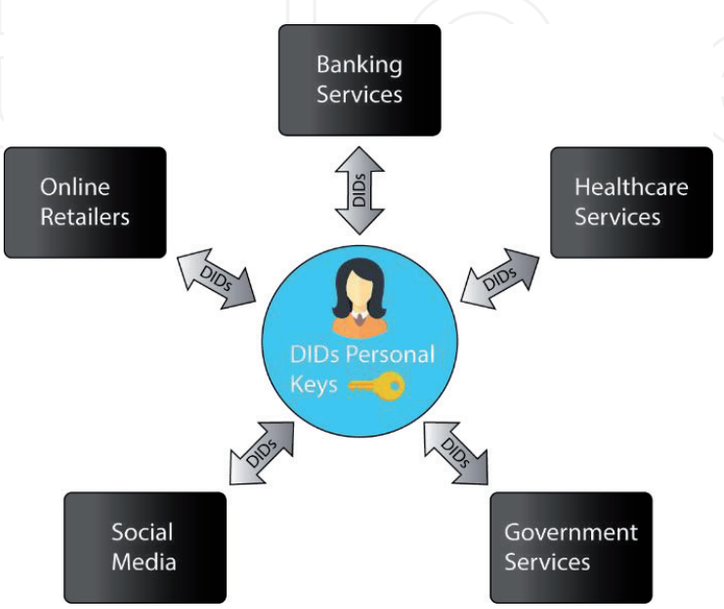

Figure 2.
Decentralized identities and service providers.

Tracking cookies and web clicks enable the linking of IDs across websites which can result in outsiders gaining a full picture of users' identity, where they live, their gender, age range, interests, and other information [10].

Figure 2 depicts how a user of several services and on-line websites can store data in a central user-controlled location and interact separately with each service provider. This enables the user to control the specific pieces of information that can be seen by each provider.

5.2 Blockchain-enabled federated identity

DIDs can help users secure and control their data property and determine who gets access to that data. Blockchains can also increase security for individuals when interacting with multiple internet platforms or services through the use of decentralized federated identities.

Blockchains allow entities to protect privacy of individuals—central to self-sovereign identity. Traditionally, users of a system or set of systems possess what is referred to as a federated identity, which can be described as a single identity used by individuals to access services or information platforms, provided by multiple parties, whereby a single identity is enabled and determined by single sign on (SSO) authentication. Consider a health care network that includes multiple entities like hospitals, insurance carriers, or urgent care clinics, where the providers enable the use of a single sign-on credential or *digital federated identity* to access all services. This type of identity, which is typically stored and managed in a central location by a service provider, is prone to security vulnerabilities [11].

The distributed nature of Blockchain technology provides an opportunity for networks to enable single sign-on, or federated identities much more securely. ElGayyar [11] proposes a blockchain-based federated identity framework (BFID) where the network of providers themselves, rather than a centralized third party, manage the system, identification, and authentication of the users. Any entity within the blockchain network can verify credentials and issue the identity for any user in the system. In a BFID, all transactions are written and maintained within the blockchain where the system takes advantage of the secure and immutable nature of the distributed ledger, thereby practically eliminating the possibility for identity breaches and potential theft.

Blockchain-based federated identity frameworks can be configured on both public and private blockchain implementations and make use of smart contracts to react to potential rule changes that may occur while governing identity management within the system. Additionally, these frameworks enable users to audit and control how their identities are used while also providing the network business entities the ability to monitor how their services are being used, enabling process improvement and a better overall user experience.

5.3 Zero-knowledge proofs

Zero-knowledge proofs enable ease of access to identity and other important data while maintaining privacy and property control for individuals. Zero-knowledge proofs are cryptographic methods whereby a user or "prover" can convince someone, or a "verifier" that something about them is true without providing, revealing or sharing that information. A common example is a customer attempting to order an alcoholic beverage from a bartender who demands to know that the patron is 21 of age or older. Providing a driver's license reveals the patron's full birth date as well as height, eye color, and home address—information that could be misused or stolen.

Zero-knowledge proofs use cryptographic algorithms that enable a prover to mathematically demonstrate to a verifier that a statement is correct without revealing any data. When the state issues a 21-and-over driver's license, it asks the driver to type in a secret nickname, unknown to the licensing bureau. This nickname could then be hashed together with the driver's license number, and stored in a public list representing valid drivers over 21. At the bar, you could type your nickname and license number into a hash generator, and if the resulting hash matched one on the list, the bartender would know that you were of legal age [12].

There are two types of zero knowledge proofs, interactive and non-interactive. Most commonly, zero knowledge protocols are interactive whereby the prover (an individual or more likely a computer) and the verifier participate in a back and forth set of questions or challenges that, when answered correctly a given number of times, enables the prover to convince the verifier, with very high probability, that the statement they are making is true.

An example of an interactive zero knowledge proof could involve two colored balls that are identical in every way accept their color. One is red and one green. Let's assume the verifier is completely color blind and cannot tell the color of either ball. You want to prove to the verifier that the balls do in fact differ in color. The verifier puts the balls behind their back and shows one. The prover indicates the color. The verifier then does this again and asks if they switched the ball. Since you can see the different colors you can say with certainty that the ball was either switched or not. After several rounds of this, it becomes more statistically true that there are in fact two different balls as the probability that you could guess correctly over and over goes down to almost zero [13].

Non-interactive proof is more like the example above of the patron proving their age to a bartender with a proof statement that reveals age but not additional information that might be revealed if the prover were to show their photo. Proving which point Value a card in a deck of 52 cards, without identifying its suit, can provide an example of this type of proof. The prover states that the card they are holding is a king but does not want to reveal which king—the king of hearts, diamonds, spades or clubs. If the cryptographic string also contains information that reveals the other 48 cards, none of which are kings, we can know for certain that the prover does in fact hold a king of some kind.

Zero knowledge proofs are powerful tools for maintaining privacy and property control for individuals that may need to provide a bit of personal information but no more than absolutely necessary.

6. Artificial intelligence and privacy

Artificial Intelligence (AI) is a broad field that includes machine learning and cognitive computing where computers are programed to mimic human cognitive functions such as learning and problem solving but many times, much faster and in more accurate ways [14]. The use of AI is expanding into a plethora of areas including speech recognition, facial recognition, medical diagnosis, financial predictions, tracking of disease outbreaks etc. AI algorithms enable computing systems to rationalize and take actions aimed at achieving a specific goal or set of goals.

User and stakeholder security can be enhanced through AI tools, which can take advantage of blockchain to open up new avenues for accessing and learning from data without taking ownership or control of that data. This can reduce risk for the organization and the stakeholders who provide the data. Both individual blockchain members and the organization or group in charge of setting governance rules and processes can benefit from building in privacy-related AI functionality (as early as possible) in the design of blockchain networks and processes.

Companies have implemented AI to create holistic views of customers by piecing together transactions from all customer touchpoints. Blockchain participants will have incentives to pull together integrated datasets by combining transactions for a single customer across all blockchain partners. This creates potential benefits for blockchain partners, but can also negatively affect the privacy of customers and other stakeholders for which this integration is possible.

In combination with options for identity protection through decentralization, AI can be used to combine personal data from blockchain participants and their stakeholders in a way that maintains information security and personal data privacy. Through these processes, user and stakeholder security can be enhanced and data sets and AI models can be improved.

We can identify four categories of stakeholders that can be affected by an organization's data transparency and privacy processes: (1) participants, whose data—both direct and indirect—are gathered; (2) victims, who are affected by decisions made using participant data; (3) users, who use participant data in their work; and (4) custodians, who manage and secure data. When AI can be used to manage access to data and to develop analytical models using that data, all stakeholders can benefit [15].

Table 1 summarizes a number of ways AI can be used in a blockchain setting to protect or increase privacy of user's personal data. This AI/Blockchain combination can increase system security by helping to detect attacks by bad actors, user security by sharing permissions and smart contracts, enable privacy-enhanced use of datasets through improved identity management and better data, and it can improve AI models through more varied, valid, and ethically-sourced data and better hypotheses. Each item in **Table 1** is described briefly below, followed by examples of use cases using this combination of technologies.

Computational intelligence (CI), a subset of AI can improve the Blockchain's attack resilience thus improving security of the system and ultimately the privacy of the data residing on the system. AI is rooted in hard computing techniques whereas computational intelligence is based on soft computing methods, which enable adaptation to a range of changing variables [16].

Computational intelligence, when combined with blockchain systems, can create more robust cryptographic functionality and ciphers thereby making it more difficult for cyber hackers to compromise systems even as computing power and efforts to hack these systems over time increases. Quite appropriately, [14] refer to the intersection of blockchain and AI as "blockchain intelligence". Additionally, AI algorithms can be built on blockchains to detect when a blockchain is under

System Security	• Malicious attack detection
	• More robust cryptographic functionality
User Security	• Users decide what data to share
	• Smart contracts can enforce established permissions
Datasets	• Improved identity masking and metadata
	• Cleaner and more accurate data
AI Models	• Broader scope and greater variety of data
	• Improved validity of data and models
	• Ethically sourced and permissioned
	• Careful construction of hypotheses

Table 1.
The role of AI in blockchain user privacy.

attack by continually monitoring blocks and activity on the chain. This technology increases trust in the system beyond what the native architecture provides [17, 18].

When blockchain participants have increased control over their own data, they have the potential to decide with which parties and for what purposes their data are shared. In order to collect participant data for use in an AI dataset, participant permissions will need to be obtained. This provides users with 'opt-in' control, rather than 'opt-out' and helps to ensure that personal data is used in ways that are consistent with the intentions of the owner. In some cases, which may become increasingly common if decentralized identity solutions are adopted, users can be compensated for providing their data to organizations seeking to utilize this data in traditional and AI decision models.

Smart contracts can also protect privacy. Permissions granted by users can be subject to complex rules embodied in smart contracts, which can enforce rules regarding the use of the data and can govern the granting and rescinding of participant data. AI can be used to scan contracts to identify participants who have or are likely to possess or provide data for desired uses.

The size and nature of datasets available in blockchain networks can also have implications for effective AI. Because many different organizations and stakeholders contribute to shared ledgers, the quantity and variability of data available for analysis can be much larger than for single-company databases. Larger datasets could enable more sophisticated identity-masking procedures, and metadata may be richer and more informative.

Because of the validation, security, timestamping and the append-only nature of blockchain ledgers, the data obtained are likely to be much cleaner and more accurate than when data are captured and maintained by many organizations in databases that are not immutable.

The ethical quality of data obtained will also be higher, and model developers and users can have increased confidence that they are following regulations. Because multi-dimensional user permissions can be granted and documented—and in some cases, enforced through smart contracts—organizations can use this data with less risk of privacy breaches. In addition, because user data can be collected using zero-knowledge proofs, complex analyses requiring specific user data can be performed and the necessary information captured and used without the need for accessing or possessing PII.

The use of blockchain data and artifacts can also result in higher quality analyses and outcomes. When data are clean and associated with clear metadata, the validity of the data is increased. Because each item in a data set is more trusted, error can be reduced and insights can be obtained through smaller data sets. When clean data are used to train AI models, those models will be more accurate, and the predictions and decisions made by those models will also be improved. Clean-training data can also be useful in validating non-blockchain data for use in AI models.

Finally, and perhaps most importantly, the hypotheses upon which AI models are built can improve, for several reasons. First, because participant permission must be obtained and possibly paid for, AI designers will need to develop clear designs that define the analyses to be performed and determine the type and amount of data needed for these analyses. This will require designers to be more aware of the universe of data that could inform these analyses and what is and is not available in distributed ledgers and personal-data files. This could help identify problems such as the lack of black faces in photo-categorizing algorithms before or during data collection.

PII may never be collected, and when it is, its use may be more intentional and usage agreements may be enforced by smart contracts. This enables more ethical approaches to gathering and managing data. AI models built using ethically-sourced and governed data can generate results that are actionable within pre-defined ethical and regulatory limits.

6.1 Emerging blockchain and AI industry uses cases

The 2019–2020 Covid 19 pandemic has prompted medical researchers and technologists to research ways to quickly gather intelligence around virus exposure and transmission as a way to combat the spread of the disease while maintaining personal privacy of users. Point-of-care diagnostics, which rely on rapid testing of patients that may have been exposed to the virus is proving to be an effective way of tracking the spread and reducing the impact of the disease. German based Pharmact AG has developed a rapid Covid 19 test that delivers results in roughly 20 minutes. This test can be used in point-of-care systems and combined with blockchain and AI to increase the speed of diagnosis and provide statistics on positive and negative results while maintaining security of personally-identifiable data. Data can be collected on blockchain infrastructure while taking advantage of the speed that AI affords to create an integrated platform that enables data from disparate sources to be analyzed. Information drawn from these systems can provide communities a powerful tool for combatting the spread of disease, reducing the burden of health care facilities and saving lives [19].

Many cities are working toward becoming "smart cities" by integrating AI and blockchain with other web 3.0 technologies such as internet of things (IoT) sensors and edge devices. Intelligent transportation systems are enabled by these technologies. Self-driving cars make use of IoT sensors to continuously monitor surrounding situations and even anticipate developments by using artificial intelligence. These cars can incorporate blockchain wallets that enable passengers to pay for rides, rentals, tolls, etc. without revealing personal data. By adding blockchain as an underlying architecture, cities and private companies can reduce the friction of renting or sharing autonomous vehicles by streamlining the process of procuring a ride. The peer to peer nature of blockchain reduces the number of people or businesses involved in the process, taking out expensive intermediaries, and reducing costs. These systems can also provide audit trails for both owners and renters, and enable rating and payment systems that maintain privacy for both parties. Data gathered by the vehicles can contribute to learning algorithms on the blockchain for increased security, scalability and efficiencies as well as improved transportation and sustainability for the city. [20]

Smart home systems that preserve user privacy while contributing usage data for analysis can likewise benefit from the integration of blockchain and AI. Smart home systems are becoming popular and manufacturers increasingly enable connectivity between devices. These systems are valuable sources of consumer usage data. AI-enabled blockchain systems can be used to push machine learning and training processes to consumer's mobile devices and edge computing servers. Users can then submit locally-trained models for analysis, in some cases with option of adding noise that makes it very difficult to trace shared data back to individual consumers. Decentralized technologies enable analysis of locally generated data without this data being submitted to a centralized server. [21]

These use cases exemplify some of the ways blockchain and AI are being used to accomplish objectives while maintaining personal data privacy. New use cases continue to be developed as technologists and user communities recognize the possibilities for systems that provide both functionality and privacy.

7. Conclusion

Blockchain and AI technologies are improving at a rapid pace and enabling possibilities for sharing and combining data in ways not previously envisioned.

At the same time, advances in these technologies provide new possibilities for the ethical use of data. Personal data, when shared, present a conundrum for firms and individuals, which can provide valuable benefits but can also create great risks and costs for both the individual and the organizations with which individual data are shared. Blockchain provides new mechanisms, such as decentralized identities and zero-knowledge proofs, that enable data to be shared in ways that maintain the privacy of the individual and allow users to maintain control over their own data. These advances can provide both increased cybersecurity and more ethical use of personal data. Blockchain participants can realize these outcomes through careful development of governance frameworks and mechanisms.

Publication of this chapter in an open access book was funded by the Portland State University Library's Open Access Fund.

Author details

Stanton Heister* and Kristi Yuthas
Portland State University, Portland, OR, USA

*Address all correspondence to: stanton.heister@pdx.edu

References

[1] RiskBased Security. 2020 Year End Report: Data Breach Quickview. [Internet]. 2021. Available from: https://pages.riskbasedsecurity.com/hubfs/Reports/2020/2020%20Year%20End%20Data%20Breach%20QuickView%20Report.pdf [Accessed: 2021-01-12]

[2] Kellerman R. Five of the Biggest Data Breaches of the 21st Century. *STAGE2DATA*. [Internet]. 2020. Available from: https://www.stage2data.com/five-of-the-biggest-data-breaches-of-the-21st-century/ [Accessed: 2020-12-17]

[3] Fruhlinger J. Equifax Data Breach FAQ: What Happened, Who was Affected, What was the Impact? [Internet]. 2020. Available from: https://www.csoonline.com/article/3444488/equifax-data-breach-faq-what-happened-who-was-affected-what-was-the-impact.html [Accessed: 2020-11-21]

[4] Statista, Figure 1. Cybersecurity Breaches and Record Exposure [Internet]. 2020. Available from: https://www.statista.com/statistics/273550/data-breaches-recorded-in-the-united-states-by-number-of-breaches-and-records-exposed/ [Accessed: 2020-11-12]

[5] Heister S, Yuthas K. Technology in Society: The Blockchain and How it Can Influence Conceptions of the Self. 2020. (60) https://doi.org/10.1016/j.techsoc.2019.101218

[6] Grimes R. What is Personally Identifiable Information (PII)? How to Protect it Under GDPR. [Internet]. 2019. Available from: https://www.csoonline.com/article/3215864/how-to-protect-personally-identifiable-information-pii-under-gdpr.html [Accessed: 2020-11-21]

[7] Uribe D, Waters G. Privacy Laws: Genomic Data and Non-Fungible Tokens. *The Journal of the British Blockchain Association*. 2020. (3) https://doi.org/10.31585/jbba-3-2-(5)2020

[8] Yoon J. Democratic Senators Introduce the Consumer Online Privacy Rights Act. [Internet]. 2019. Available from: https://www.insideprivacy.com/united-states/congress/democratic-senators-introduce-the-consumer-online-privacy-rights-act/ [Accessed: 2021-01-08]

[9] Consensys. Busting the Myth of Private Blockchains. [Internet]. 2020. Available from: https://consensys.net/enterprise-ethereum/best-blockchain-for-business/busting-the-myth-of-private-blockchains/ [Accessed: 2020-12-12]

[10] Microsoft. Decentralized Identity. [Internet]. Available from: https://www.microsoft.com/en-us/security/business/identity/own-your-identity [Accessed: 2020-12-12]

[11] ElGayyar M, ElYamany H, Grolinger K, Capretz M, Mir S. Blockchain-Based Federated Identity and Auditing. *International Journal of Blockchains and Cryptocurrencies.* 2020.p. 179-205

[12] Lesavre L, Varin P, Mell P, Davidson M, Shook J. A Taxonomic Approach to Understanding Emerging Blockchain Identity Management Systems. [Internet]. 2019. *COMPUTER SECURITY RESOURCE CENTER.* Available from: https://csrc.nist.gov/publications/detail/white-paper/2019/07/09/a-taxonomic-approach-to-understanding-emerging-blockchain-idms/draft [Accessed: 2020-11-17]

[13] Wikipedia. Zero-Knowledge Proof. *Wikipedia.* [Internet]. 2020. Available from: https://en.wikipedia.org/wiki/

Zero-knowledge_proof [Accessed: 2021-01-18]

[14] Zheng Z, Dai H, Wu J. Blockchain Intelligence: When Blockchain Meets Artificial Intelligence 2020. *arXiv preprint arXiv:1912.06485*.

[15] Bertino E, Kundu A, Sura Z. Data Transparency with Blockchain and AI Ethics. Data and Information Quality; 2019. https://doi.org/10.1145/3312750

[16] Wikipedia. Computational intelligence. *Wikipedia.* [Internet]. Available from: https://en.wikipedia.org/wiki/Computational_intelligence#:~:text=According%20to%20Bezdek%20(1994)%2C,a%20subset%20of%20Artificial%20Intelligence.&text=Crisp%20logic%20is%20a%20part,be%20partially%20in%20a%20set. [Accessed: 2021-01-18]

[17] Marwala T, Xing B. Blockchain and Artificial Intelligence 2018. Arvix preprint arXiv:1802.04451.

[18] Salah K, Rehman M, Nizamuddin N, Al-Fuqaha A. Blockchain for AI: Review and Open Research IEEE; 2019. (7) p.10127-10149. DOI: 10.1109/ACCESS.2018.2890507

[19] Mashamba-Thompson T, Crayton E. Self-Testing: Blockchain and Artificial Intelligence Technology for Novel Coronavirus Disease 2019. Diagnostics. 2020. (10) 198. https://doi.org/10.3390/diagnostics10040198

[20] Singh S, Sharma P, Yoon B, Shojafar M, Cho G, Ra I. Convergence of Blockchain and Artificial Intelligence in IoT Network for the Sustainable Smart City. 2020. *Sustainable Cities and Society*. (63) art. no. 102364

[21] Zhao Y, Zhao J, Jiang L, Tan R, Niyato D, Li A, Lyu L, Liu Y. Privacy-Preserving Blockchain-Based Federated Learning for IoT Devices. IEEE Internet of Things Journal. 2020.