Blockchain-based data governance for privacy-preserving in multi-stakeholder settings

Rodrigo Dutra Garcia [1] **Jó Ueyama** [1] **(Advisor)**

[1]Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo
Avenida Trabalhador São-Carlense 400, São Carlos, SP, 13566-590 – Brazil

{rgarcia, joueyama@icmc.}@usp.br

Abstract. *In multi-stakeholder systems, such as healthcare, the Internet of Things, and supply chain management, there is frequent data generation, exchange, and sharing. As a result, data owners often desire control over their data and maintain privacy, while data consumers require methods to ascertain the origins and creators of the data. These conflicts of interest require developing data governance systems that guarantee data provenance, privacy protection, consent management, and selective disclosure. This research proposed a decentralized data governance system utilizing blockchain technology, proxy re-encryption (PRE), and Boneh, Boyen, and Shacham (BBS) signatures to address these challenges. The proposed system enables data owners to control, selectively share, and track their data through privacy-enhancing, consent management, and selective disclosure mechanisms while also allowing data consumers to understand the lineage of the data through a blockchain-based provenance mechanism. As a case study, the research examined and evaluated electronic prescriptions involving sensitive data and multiple stakeholders, including patients as data owners and doctors and pharmacists as data consumers. The research was structured as a collection of published articles organized in the following sequence: problem formulation and developing smart contracts, implementing privacy and consent management through PRE, and applying BBS signatures for selective data sharing. The proof-of-concept implementation and evaluations, conducted using CosmWasm, Hyperledger Besu, Ethereum, pyUmbral PRE, and BBS signatures, demonstrate that the proposed decentralized system is platform-agnostic, scalable, and capable of providing a higher level of transparency, privacy, and trust with minimal overhead.*

1. Introduction

The expansion of digital technologies has led to exponential data production and sharing among stakeholders. Applications such as healthcare and the Internet of Things are examples of this trend, as they rely heavily on collecting and disseminating data [Mukta et al. 2022]. However, as the scale of data sharing expands, it is crucial to ensure that individuals' privacy rights are protected. In multi-stakeholder applications, data consumers need to clearly understand the lineage of the data they utilize, including knowledge of the services and companies involved in collecting, storing, and disseminating the data. Furthermore, data owners need to consent to share specific information with data consumers and have the control to grant or revoke access to personal

information and sensitive data, allowing greater control and transparency in handling personal information [Kakarlapudi and Mahmoud 2021].

In the healthcare industry, sensitive data such as electronic medical records (EMRs) are routinely produced and shared among various stakeholders, including patients, doctors, hospitals, and pharmacies. For instance, EMRs contain personally identifiable information (PII), diagnosis, and medication. This information is required to provide quality care. However, handling sensitive data requires robust data protection measures to ensure the privacy and security of individuals. The centralization of data storage in current healthcare systems, including electronic prescription systems, imposes a significant challenge to protecting EMRs and the privacy of individuals whose information is being collected and shared. The centralized structure of these systems creates a single point of failure, making them vulnerable to breaches and unauthorized access [Wazid et al. 2022]. Additionally, the lack of transparency inherent in such architecture can compromise the ability of patients to exert control and oversight over their personal information, raising concerns about potential violations of privacy rights [Qahtan et al. 2023].

The trust mechanism in a centralized architecture is typically based on a central authority to enable controlling and managing data access. In contrast, a decentralized architecture, such as blockchain, uses distributed ledger technology to record and share information in a tamper-proof manner without needing a central entity [Nakamoto 2009]. Blockchain technology uses a peer-to-peer (P2P) network to establish trust among participating nodes through a consensus mechanism rather than relying on a central authority, providing security and transparency. The transparency inherent in blockchain technology enables all participants in the network to have a clear view of the data and its history, making it easier to trace the lineage of the data and understand how it has been shared and processed. Furthermore, the decentralized structure of the network makes it more resilient to attacks and failures. In addition to its decentralized structure, blockchain technology also enables smart contracts. This feature was initially proposed by Szabo [Szabo 1997] as protocols, and in blockchain platforms such as Ethereum, acts as immutable self-executing programs written in code and stored on the blockchain. It enables task automation and agreements without needing a third party as an intermediary, which increases efficiency, security, and trust in the execution of the contract while also reducing costs [Hewa et al. 2021].

Despite capabilities, one of the main limitations of using blockchain technology in sensitive applications such as healthcare and the Internet of Things is the issue of data privacy [Peng et al. 2021]. The transparency of blockchain technology means that all data stored on the blockchain is available to all network nodes. It implies a significant challenge to maintaining the confidentiality of sensitive healthcare information, such as medical records and personal information, which must be protected in compliance with regulations such as the General Data Protection Regulation (GDPR). Researchers have proposed some approaches to address the issue of sensitive data privacy in healthcare applications, such as off-chain data storage, encryption, and zero-knowledge proofs (ZKP) [Yin et al. 2023]. However, there is a lack of study on enabling data owners to manage and selectively share attributes with stakeholders while preserving sensitive data privacy, particularly in the healthcare sector, such as electronic prescription.

2. Research Questions

The objective of the research is to propose a blockchain-based system that answers the following research questions (RQ):

- **RQ1:** how can blockchain and smart contracts secure and manage sensitive data in a tamper-proof ledger? In particular, how can smart contracts be implemented for electronic prescription use cases using byzantine fault tolerance (BFT) platforms such as Tendermint?
- **RQ2:** how can data owners, such as patients, maintain their privacy while still tracking and governing the usage by other parties?
- **RQ3:** while maintaining data owners' privacy, how can a regulatory entity access data for accountability and compliance verification in a decentralized data governance system?
- **RQ4:** how can data owners selectively share specific attributes with certain stakeholders in a reliable and scalable manner?

Particularly, this study employed proxy re-encryption (PRE) to ensure the privacy of sensitive data and enable data sharing with owner consent. Additionally, Boneh, Boyen, and Shachum (BBS) signatures built with zero-knowledge proof were utilized to allow the selective sharing of data in a blockchain-based system. The research focused on the electronic prescription (e-prescription) use case, a multi-stakeholder application with sensitive data sharing [Vejdani et al. 2022]. Patients act as data owners and can selectively share their data with relevant stakeholders, such as pharmacies and doctors, while keeping the data securely encrypted and stored on the blockchain [Garcia 2023].

3. Publications and Contributions

The dissertation was structured as a series of published articles (four articles in total) grouped according to their contribution. Figure 1 shows the research questions, techniques used, and published articles. It also includes the publication metrics over the period from 2017 to 2020, as evaluated using the Qualis Capes platform[1].

3.1. Towards a Decentralized e-Prescription System Using Smart Contracts

The first article, titled *"Towards a Decentralized e-Prescription System Using Smart Contracts"* [Garcia et al. 2021], introduces an electronic prescription system that leverages smart contracts on BFT platforms. The main contributions of this research are as follows:

- Design and implementation of a blockchain-based e-prescription system using smart contracts on a BFT-based consensus mechanism. In particular, the work used Tendermint consensus, which is not widely adopted for smart contract applications.
- Performance evaluation of the proposed system with another BFT platform, Hyperledger Fabric, evaluating contract file size, transaction overhead, scalability, and smart contract deployment complexity.

However, the study does not evaluate and compare the cost with another existing consensus mechanism, such as Proof of Work (PoW). Furthermore, the work does not address the issue of data privacy in transactions and the mechanisms used to protect it.

[1] https://qualis.capes.gov.br/

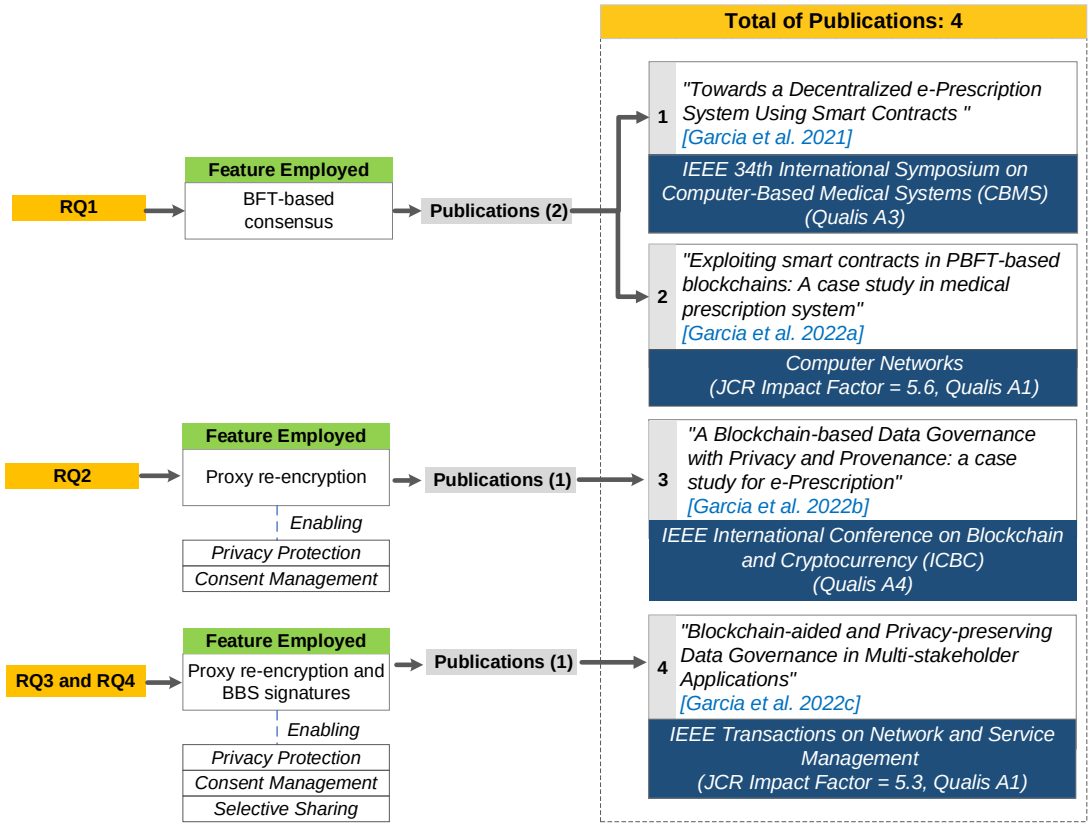

Figure 1. The research questions highlighting the features employed and the published articles during the research

3.2. Exploiting smart contracts in PBFT-based blockchains: A case study in medical prescription system

The second article, titled *"Exploiting smart contracts in PBFT-based blockchains: A case study in medical prescription system"* [Garcia et al. 2022a] builds upon the previous work by providing a more comprehensive evaluation and discussion. The contributions are the follows:

- Evaluate the implementation of smart contracts on BFT blockchain platforms such as Tendermint and Hyperledger Besu;
- Compare their operational cost and performance to Ethereum, a PoW blockchain.

However, the study only briefly discusses the use of public key encryption to preserve patient privacy in the blockchain-based e-prescription system and does not evaluate its effectiveness.

3.3. A Blockchain-based Data Governance with Privacy and Provenance: a case study for e-Prescription

The third work, entitled *"A Blockchain-based Data Governance with Privacy and Provenance: a case study for e-Prescription"* [Garcia et al. 2022b], implements the use of proxy re-encryption to ensure patient consent and privacy. The main contribution of the research is:

- A system that enables data owners to control and monitor their data through privacy-enhancing and consent management mechanisms while also allowing data consumers to trace the origins of the data through a blockchain-based provenance system.

However, the study does not implement a selective disclosure mechanism to allow data owners to share specific information with selected stakeholders.

3.4. Blockchain-aided and Privacy-preserving Data Governance in Multi-stakeholder Applications

As an extension of prior work, the fourth article, entitled *"Blockchain-aided and Privacy-preserving Data Governance in Multi-stakeholder Applications"* [Garcia et al. 2022c], includes a regulator authority and incorporates the use of Boneh, Boyen, and Shacham signatures to enable selective sharing by data owners. The research presents the following contributions:

- A decentralized architecture for multi-stakeholder applications that combines blockchain, smart contracts, and proxy re-encryption mechanism that allows for data owner consent while keeping data encrypted on the blockchain, enabling regulator entity to track the records with data owner permission;
- The use of BBS signatures to enable data owners to share specific attributes with data consumers;
- A proof-of-concept performance evaluation using BFT blockchain platforms (CosmWasm and Hyperledger Besu) compared to Ethereum PoW. Additionally, the study employs NuCypher's pyUmbral proxy re-encryption (PRE) library and MATTR JSON-LD library using BLS12-381 key pairs for BBS signature evaluation.

All evaluation software and smart contracts developed as part of the research can be accessed on GitHub[2].

4. Discussion

This section discusses the challenges and importance of results associated with implementing a blockchain governance system.

4.1. Challenges

There are some challenges associated with implementing a blockchain governance system. For instance, regulatory compliance is crucial, as blockchain systems must ensure strict adherence to privacy regulations and rigorous data protection standards such as the GDPR and LGPD (General Personal Data Protection Law). Another challenge includes the integration of blockchain-based governance into existing infrastructures. The existing reliance on centralized systems presents both technical and operational challenges. Additionally, from an adoption standpoint, scalability is a challenge; the technology must demonstrate that it can handle large transaction volumes while maintaining privacy—an essential factor in complex, multi-stakeholder environments, including healthcare, supply chain, and the Internet of Things.

[2]`https://github.com/rodrigodg1/e-prescription`

4.2. The BBS Signatures Employment

This work uses BBS signatures, a cryptographic technique for privacy-sensitive data sharing. Data owners can choose which specific pieces of information to reveal to different parties. While other methods like Redactable Signatures exist [Hörandner et al. 2020], they are based on less robust security assumptions. BBS signatures support Zero-Knowledge Proofs. The security of BBS is based on the robust q-SDH (Strong Diffie-Hellman) assumption, making it stronger and more secure against cryptographic attacks [Yamamoto et al. 2022].

4.3. Importance of the Results

The proposed system prioritizes privacy and data ownership. It empowers data owners, like patients and users, to directly control who accesses their information – a vital feature in today's world of frequent data breaches and heightened privacy concerns. Blockchain technology guarantees transparency and traceability in how data is used and where it comes from. This builds trust in environments with many stakeholders and provides clear audit trails for regulatory compliance. Additionally, smart contracts automate processes, save costs, and significantly improve operational efficiency, especially in sectors with data-sharing requirements. This research advances the knowledge of combining advanced cryptographic techniques with blockchain technology to address practical data governance challenges.

5. Conclusion

Applications that involve multiple stakeholders, such as healthcare, the Internet of Things, and supply chain management, necessitate the protection of privacy and the management of private data. This research presents a governance system based on blockchain technology, which explores e-prescription to ensure privacy, consent management, and selective sharing. The research is comprised of a collection of published articles that contribute to the field in the following order: the implementation and evaluation of smart contracts using BFT-based platforms like Tendermint consensus, Hyperledger Fabric, and Hyperledger Besu, with a comparison of operational costs to Ethereum PoW; the implementation and evaluation of proxy re-encryption operations in the context of e-prescription, utilizing the NuCypher pyUmbral PRE library for enabling patient consent; and the implementation and evaluation of BBS signatures, which enable selective sharing by the patient through the use of the MATTR JSON-LD library.

To address the first research question (RQ1) on how blockchain and smart contracts can be employed to secure and manage sensitive data in an immutable ledger, the first article proposes a decentralized e-prescription system that utilizes smart contracts. The proposed system aims to securely manage electronic medical records by leveraging the tamper-proof and transparent features of blockchain, thereby mitigating fraudulent activities such as data tampering. Furthermore, the contracts are implemented, evaluated, and compared on two BFT-based platforms: CosmWasm (Tendermint consensus) and Hyperledger Fabric. The findings indicate that a higher number of validator nodes enhances the system's fault tolerance. However, it also leads to increased latency due to BFT consensus message traffic. Additionally, the second article expands on the evaluation and privacy considerations within multi-stakeholder applications like e-prescription. Specifically, the study compares the operational costs of smart contracts in

CosmWasm and Hyperledger Besu, focusing on the CPU and memory usage of validator nodes during consensus. Moreover, the article evaluates the block time of BFT platforms in comparison to Ethereum mining time. The results demonstrate the feasibility of BFT-based platforms for healthcare applications in contrast to PoW solutions.

In conjunction with the second article, the third article presents a blockchain-based system that addresses the privacy protection and consent management requirements of multi-stakeholder environments. The system aims to respond to the second research question (RQ2), which pertains to preserving the privacy of data owners and enabling them to track and govern the usage of their data by other parties. The system ensures the encryption of all data and stores requests in a transparent and tamper-proof ledger, allowing data owners to delegate access to other parties. Notably, the study employs proxy re-encryption as a means to safeguard data privacy and enable data owner consent in data sharing. The evaluations demonstrate that proxy re-encryption effectively protects the privacy of patients and enables data governance with minimal overhead.

Furthermore, to address RQ3 concerning how can a regulatory entity access data for accountability and compliance verification in a decentralized data governance system, the fourth article, using patient permission, includes smart contracts to enable the regulator entity to track the flow of goods through a tamper-proof ledger reducing illegal drug sales. Regarding RQ4 about how data owners can selectively share specific attributes with certain stakeholders in a reliable and scalable manner, the fourth article includes BBS signatures on top of proxy re-encryption to allow data owners to selectively share data while maintaining the encrypted form on the blockchain. In particular, the patient acts as a data owner and creates a verifiable derived proof using zero-knowledge proof to share specific attributes with selected stakeholders. The evaluations show that the blockchain-based governance system using proxy re-encryption and BBS signatures libraries can be explored in any multi-stakeholder system with private data with minimal overhead.

Acknowledgments

The author would like to express sincere thanks to advisor Professor Dr. Jó Ueyama for his support during the various stages of this research. Additionally, the author extends deep gratitude to Dr. Gowri Ramachandran for his suggestions and contributions to improving this research. The author is also grateful to the Brazilian National Council for Scientific and Technological Development (CNPq) for financial support during this research (Grant No. 133470/2020-2).

References

Garcia, R. D. (2023). *Blockchain-based data governance for privacy-preserving in multi-stakeholder settings*. PhD thesis, Universidade de São Paulo. Agência de Bibliotecas e Coleções Digitais.

Garcia, R. D., Ramachandran, G., and Ueyama, J. (2022a). Exploiting smart contracts in PBFT-based blockchains: A case study in medical prescription system. *Computer Networks*, page 109003.

Garcia, R. D., Ramachandran, G. S., Jurdak, R., and Ueyama, J. (2022b). A Blockchain-based Data Governance with Privacy and Provenance: a case study for e-Prescription.

2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 00:1–5.

Garcia, R. D., Ramachandran, G. S., Jurdak, R., and Ueyama, J. (2022c). Blockchain-aided and Privacy-preserving Data Governance in Multi-stakeholder Applications. *IEEE Transactions on Network and Service Management*, PP(99):1–1.

Garcia, R. D., Zutião, G. A., Ramachandran, G., and Ueyama, J. (2021). Towards a decentralized e-prescription system using smart contracts. *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, 00:556–561.

Hewa, T., Ylianttila, M., and Liyanage, M. (2021). Survey on blockchain based smart contracts: Applications, opportunities and challenges. *Journal of Network and Computer Applications*, 177:102857.

Hörandner, F., Ramacher, S., and Roth, S. (2020). Selective end-to-end data-sharing in the cloud. *Journal of Banking and Financial Technology*, 4(1):139–157.

Kakarlapudi, P. V. and Mahmoud, Q. H. (2021). A systematic review of blockchain for consent management. *Healthcare*, 9(2).

Mukta, R., young Paik, H., Lu, Q., and Kanhere, S. S. (2022). A survey of data minimisation techniques in blockchain-based healthcare. *Computer Networks*, 205:108766.

Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system.

Peng, L., Feng, W., Yan, Z., Li, Y., Zhou, X., and Shimizu, S. (2021). Privacy preservation in permissionless blockchain: A survey. *Digital Communications and Networks*, 7(3):295–307.

Qahtan, S., Yatim, K., Zulzalil, H., Osman, M. H., Zaidan, A., and Alsattar, H. (2023). Review of healthcare industry 4.0 application-based blockchain in terms of security and privacy development attributes: Comprehensive taxonomy, open issues and challenges and recommended solution. *Journal of Network and Computer Applications*, 209:103529.

Szabo, N. (1997). Formalizing and securing relationships on public networks. *First monday*.

Vejdani, M., Varmaghani, M., Meraji, M., Jamali, J., Hooshmand, E., and Vafaee-Najar, A. (2022). Electronic prescription system requirements: a scoping review. *BMC Medical Informatics and Decision Making*, 22(1):1–13.

Wazid, M., Das, A. K., Mohd, N., and Park, Y. (2022). Healthcare 5.0 Security Framework: Applications, Issues and Future Research Directions. *IEEE Access*, 10:129429–129442.

Yamamoto, D., Suga, Y., and Sako, K. (2022). Formalising linked-data based verifiable credentials for selective disclosure. In *2022 IEEE European Symposium on Security and Privacy Workshops (EuroSPW)*, pages 52–65.

Yin, R., Yan, Z., Liang, X., Xie, H., and Wan, Z. (2023). A survey on privacy preservation techniques for blockchain interoperability. *Journal of Systems Architecture*, 140:102892.