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

Rodrigo Dutra Garcia*, Gowri Sankar Ramachandran†, Raja Jurdak†, and Jo Ueyama*
*Institute of Mathematics and Computer Science, University of São Paulo, Brazil. {rgarcia,joueyama@icmc.}@usp.br
†School of Computer Science, Queensland University of Technology, Australia. {g.ramachandran,r.jurdak}@qut.edu.au

2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC) | 978-1-6654-9538-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICBC54727.2022.9805545

***Abstract*—Real-world applications in healthcare and supply chain domains produce, exchange, and share data in a multi-stakeholder environment. Data owners want to control their data and privacy in such settings. On the other hand, data consumers demand methods to understand when, how, and who produced the data. These requirements necessitate data governance frameworks that guarantee data provenance, privacy protection, and consent management. We introduce a decentralized data governance framework based on blockchain technology and proxy re-encryption to let data owners control and track their data through privacy-enhancing and consent management mechanisms. Besides, our framework allows the data consumers to understand data lineage through a blockchain-based provenance mechanism. We have used Digital e-prescription as the use case since it has multiple stakeholders and sensitive data while enabling the medical fraternity to manage patients' prescription data, involving patients as data owners, doctors, and pharmacists as data consumers. Our proof-of-concept implementation and evaluation results based on CosmWasm and pyUmbral PRE show that the proposed decentralized system guarantees transparency, privacy, and trust with minimal overhead.**

***Index Terms*—Data Governance, Decentralized, E-prescription, Privacy, Blockchain, Smart Contracts, Proxy Re-encryption**

I. Introduction

Prescription systems allow healthcare professionals, such as physicians, to create digital records about a patient's health status by adding diagnosis and medications data. It allows for more efficient communication and reduced inconsistencies compared to paper-based prescriptions [1, 2]. Thus, digital prescription systems have increased globally, enabling multiple stakeholders, including doctors and pharmacies, to effectively access and manage patients' data.

Patients want to control their data and privacy in healthcare settings since prescription and diagnosis data contain sensitive and personally identifiable information. Note that unauthorized parties may gain access and misuse patients' data [3]. Therefore, it is essential to *protect patients' privacy while letting them manage and permit access to their data transparently*, which is one of the problems this paper aims to investigate.

Pharmacies must sell certain drugs such as antibiotics with a valid doctor's prescription. A prescription containing an antibiotic medicine is valid for only a single purchase, meaning the pharmacy and the patient must obey the recommended dosages. However, pharmacies tend to sell medications illegally to patients to gain financial revenue, even with the old and used prescription. Such illegal sales would lead to unwanted side effects, including drug abuse and overdoses [4, 5], burdening the healthcare system. Therefore, it is essential to regulate the medicine supply chain to prevent the unauthorized sales of medications, which is one of the focuses of this work.

Existing digital prescription systems primarily employ a centralized architecture, offering limited to no visibility into the operations providing maximum power to the administrating organization [1, 2]. Such centralized architectures are susceptible to single points of failure, enabling opportunities for data tampering. In addition, centralized systems may also misuse patients' health data without their consent, resulting in privacy violations. In summary, centralized architectures offer no transparency undermining the integrity of medical information while affecting patients' privacy [6]. We, therefore, argue that a decentralized architecture with support for consent management, privacy preservation, and data provenance is essential for a trusted digital prescription system.

Existing works in digital e-prescription do not securely manage consent while providing support for privacy protection and accountability [7, 8, 9]. We propose a decentralized data governance framework for the electronic prescription that:

- Helps patients *store, manage, and share* prescription data with other stakeholders through a tamper-proof ledger.
- Protects patients' *privacy* by storing encrypted prescription data on the blockchain ledger to withhold personally identifiable and sensitive information from third parties, including drug regulators.
- Provides support for *consent management* using proxy re-encryption scheme and smart contracts.
- Supports *data provenance* to let data owners and data consumers efficiently monitor the historical records of the data and its origin, including who accessed the data and for what purposes.
- Enables the drug regulators to control and monitor the flow of medications to the pharmacies through the *accountable* blockchain ledger, thereby limiting illegal sales.

We have developed a proof-of-concept implementation using the CosmWasm, which uses Tendermint (a Byzantine Fault Tolerance (BFT) consensus mechanism) and NuCypher pyUmbral [10] proxy re-encryption (PRE) library to estimate the overhead and feasibility. Our evaluation results show that the proposed data governance framework introduces minimal

overhead while letting data owners control and manage their data with transparency and trust guarantees. Although we discuss the data governance framework through an e-prescription use case, the proposed framework is suitable for any multi-stakeholder application, including supply chain management, dealing with digital and sensitive data.

II. Related Work

Electronic prescription systems operate in a multi-stakeholder environment. It requires the integrity and transparency of information to avoid illegal drug sales while preventing patients' health problems due to drug overdose. Besides, the application of privacy-preserving techniques for medical records is another requirement to avoid the misuse of sensitive information present in prescriptions. Alnafrani and Acharya proposed SecureRx [7], a blockchain solution using the Ethereum platform to maintain patient records and prescriptions. Garcia et al. [8] proposed a decentralized e-prescription system using smart-contracts on a BFT platform. However, these solutions do not *manage consent* and focus on writing records to an immutable ledger without providing mechanisms to *track who accessed the data and for what purposes* while *protecting patient's sensitive information*.

Other research works investigate approaches to ensure the integrity and privacy of medical records by preventing tampering and data leakage. Zou et al. proposes SPchain [11], a blockchain and PRE-based solution for sharing electronic health records (EHR). Li et al. introduced DMMS [9], a solution that exploits blockchain technology for medication history management and electronic prescriptions. Bhaskaran et al. [12] proposed a solution to store consumers' data in encrypted form on blockchain. Entities that wish to get access can raise consent requests on the chain, and the data owners can provide such consent cryptographically. However, the works above do not manage patient consent for sharing sensitive data between multi-stakeholder applications using the PRE mechanism in the electronic prescriptions use case.

III. Background on Proxy Re-Encryption (PRE)

Proxy re-encryption is an asymmetric encryption technique initially proposed by Blaze et al. [13] in which an entity A (delegator) can delegate the decryption rights to another entity B (delegatee) through a proxy server.

Initially, a message m is encrypted using the delegator's public key, $C_A = Enc(pk_A, m)$, and stored in a database. If delegatee B needs to decrypt the message, he must initially request decryption rights for the delegator, informing his public key pk_B. If the delegator agrees, it will produce a delegation key $rk_{A\to B}$ and send it to the proxy.

For delegatee B to be able to decrypt the information, the proxy server must use the delegation key $rk_{A\to B}$ to re-encrypt C_A, that is, $C_B = ReEnc(rk_{A\to B}, C_A)$. After re-encryption, the delegatee can use his private (i.e., secret) key sk_B and decrypt the message. At all stages, only the public key is shared between the participants. From the proxy's point of view, it does not learn or try to decrypt confidential information. It receives encrypted information C_A and sends other encrypted information C_B.

IV. Decentralized Architecture with Consent Management and Privacy Protection

A. System Model and Threats

We assume a system comprising of patients, doctors, and pharmacies. When a patient visits a doctor, the doctor creates a new medical record that includes diagnostic data, personal details such as name and age, and prescriptions.

We focus on the following threats:

- Privacy threat: The patient's medical record includes sensitive data, which should not be revealed to unauthorized third parties without the patient's consent.
- Illegal drug sales: The lack of visibility into the medication supply chain leads to illegal medication sales, resulting in drug overdoses.

Given these threats, this work aims to develop a solution with the following objectives:

Objective 1: We aim to develop a transparent medical prescription system based on the blockchain without revealing sensitive data to unauthorized third parties. Note that the data stored on the blockchain is visible to the public on a blockchain platform. *Can we allow the patients to store and manage medical data in a tamper-proof ledger without violating patients' privacy?*

Objective 2: When the data get stored on a digital system, doctors and health care agencies can access the data for diagnostic and survey purposes. Under this circumstance, it is important to let patients or data owners have visibility into data usage. *Can we allow the data owners to track and govern data usage by other parties?*

B. Proposed Solution

We propose a decentralized data governance framework using blockchain technology and smart contracts to help patients manage their data more efficiently. When a patient consults a doctor, the doctor creates prescription data by recording the diagnosis, recommended medications, and dosage. Then, the prescription data is encrypted using the patient's public key and stored in the blockchain via the smart contract. We assume that the patient shares her public key with the doctor when she makes an appointment to see the doctor. The patient needs to allow the pharmacists access to the prescription data to receive medication from the pharmacy. We propose a data access tracking mechanism within the contract state to monitor data accesses. In this way, any query or update in the status of the records will be registered. Note that the existing blockchain-based systems support writing data to an immutable ledger. Still, they do not monitor or provide support for governing data usage, which is necessary for data provenance and privacy. Our framework not only records the data on an immutable ledger in a privacy-preserving manner but also logs access requests to govern data usage.

To prevent illegal medications sales by the pharmacy, the regulatory agency will count, through a control contract, the

number of drugs supplied to the pharmacy with the number of drugs sold (in the sales contract). We assume that the blockchain can hold encrypted prescription data for brevity. Patients can report the pharmacy that sells unlawful drugs and receive rewards (tokens) through a smart contract. We can extend the framework by storing the encrypted prescription data on off-chain storage while maintaining the hash on-chain, which we plan to tackle in our future work.

C. Proxy Re-Encryption Mechanism

We use the proxy re-encryption technique to ensure data privacy in this work. In this way, stakeholders can decrypt prescription data only with the patient's consent via the delegation mechanism. The diagram in Figure 1 shows the architecture with the PRE operations:

1) From appointment with the patient, the doctor creates a prescription containing the items: personal information (*PI*), medication (*MED*), and diagnosis (*DIA*) for future analysis. Before sending the prescription to the $create\_prescription$ smart contract method and being stored in the contract state, the prescription items are encrypted by the doctor application separately using the patient's public key (pk_P). *Therefore, the patient has flexibility and can consent to data sharing*.
2) For the doctor, pharmacy, or regulator to analyze any item in the prescription, it will be necessary to request decryption rights from the patient, informing their respective public key (pk).
3) If the patient agrees with the request, the patient's application will generate a delegation key. The delegation key will be encrypted and sent to the *set_consent* contract method.
4) In the stakeholder application, the proxy will perform the re-encryption (RE) operation using the respective delegation key and the allowed prescription item. The doctor has access to all the prescription data. For this, the proxy will perform the re-encryption for personal information C_{PI}, medication C_{MED} and diagnosis C_{DIA}. The pharmacy and the regulator can only re-encrypt the prescribed medication.
5) After the re-encryption step, stakeholders can decrypt the item with their respective private key (s_k) and analyze the information allowed by the patient.

Note that sensitive prescription items are encrypted before being stored in the blockchain. In this way, records are private and immutable. Other organizations will only be able to decrypt the information with the patient's permission. The proxy re-encryption mechanism is a privacy software module implemented in the stakeholder application to act on confidential data sharing operations.

Note about proxy: a proxy is a software that only re-encrypts information. The proxies do not store any private keys and do not see any message from the ciphertexts. From their perspective, they only see an incoming ciphertext and the result after re-encryption, which is also a ciphertext.

Consent mechanism and delegation key: Figure 2 represents steps 2 and 3 of Figure 1 where each stakeholder is a full node (i.e., containing the PRE operations and blockchain). In step 2, the stakeholder sends a request to the patient through a consent contract. In step 3, the patient will create and encrypt the delegation key using the stakeholder's public key. In this way, requests are transparent to all network participants, and only the stakeholder can decrypt the delegation key.

D. How does the proposed solution meet the objectives?

Objective 1 focuses on providing transparency to the prescription system while protecting patients' privacy — our solution stores the patients' data on the blockchain but in an encrypted form. The prescription data is made available to other parties after the patient's consent.

Objective 2 focuses on governing data usage - our solution tracks the data access requests of consumers and permissions of data owners through a smart contract and immutable ledger. Therefore, data owners can have visibility into their data and its usage. We understand that a malicious data consumer may access the patient's data with their permission and then post it on a black market or other digital platforms. We plan to investigate digital watermarking and steganography in our future work to overcome this problem [14].

V. Proof-of-Concept Implementation and Evaluation

A. Privacy: Proxy Re-Encryption

Evaluation Goals: To understand the overhead and feasibility of PRE operations, we evaluate execution time tests for steps in Figure 1. We evaluated encrypting (step 1), creating a delegation key (step 3), re-encryption (step 4), and decrypting (step 5).

Evaluation Setup and Methodology: The PRE evaluation programs and scripts were implemented in Python programming language using NuCypher pyUmbral PRE technology, an open-source implementation that uses the *secp256k1* elliptic curve [15]. We use the *time* module to calculate the difference between the start and end of each operation. All software created for evaluation is available on GitHub [16].

To identify realistic file sizes for medications and dosage prescriptions, we used the English Prescribing Dataset [17]. Prescription items used for evaluation are represented in separate text files with different sizes ranging from 0.43 Kilobyte (kB) to 0.82 kB for personal information, 0.24 kB to 0.53 kB for medication, and 2.18 kB to 8975.74 kB $\approx$ 8.76 Megabyte (MB) for diagnosis. While the file sizes are inferred from [17], file contents are randomly generated by the evaluation software. In total, 1000 iterations were performed for different file sizes. We used a Linux virtual machine with an Intel Core i7-10510U 1.80GHz (Dual-Core) processor and 6GB of RAM for the evaluation.

Execution Time Evaluation: Figure 3 shows the average execution time for application-level PRE operations using the data files.

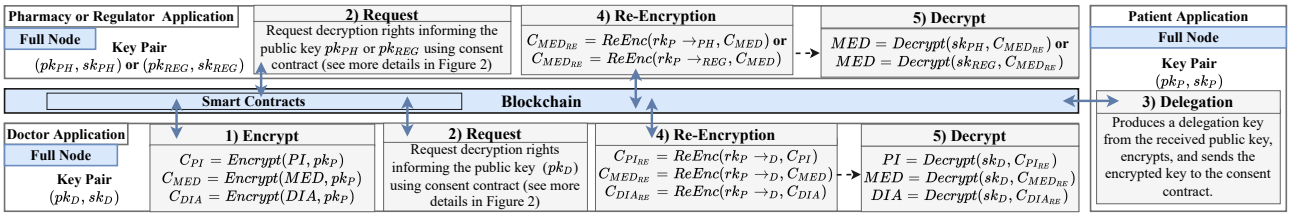

Fig. 1. Our decentralized data governance framework with support for PRE mechanism, consent management, data provenance, and privacy.

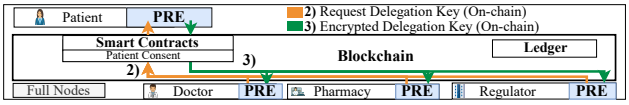

Fig. 2. On-chain request mechanism (step 2) and sending the encrypted delegation key (step 3) using smart contracts

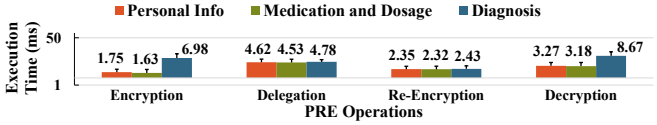

Fig. 3. Average execution time in PRE operations

To encrypt the diagnostic data (step 1), it took an average of 6.98 milliseconds (ms), while medication and personal information data took an average of 1.63 ms and 1.75 ms, respectively. There were slight variations in the average processing times in the delegation and re-encryption stage for the prescription items. The delegation stage (step 3) took an average of around 4 ms, while in the re-encryption operation (step 4), the average execution time was around 2 ms for all prescription data. In the decrypt stage (step 5), the item that obtained the highest execution average was the diagnosis with 8.67 ms. In comparison, medication and the patient's personal information took an average of around 3 ms.

These results show PRE operations' execution time cost is relative to the data size. In our evaluation, even with text files with sizes in Megabyte, the operations did not exceed 50 ms to be executed. In this sense, our proposed framework protects the privacy and manages consent with a low operational overhead. We also believe PRE operations can run on platforms like Raspberry Pi or mobile phones.

B. Smart Contract: CosmWasm Implementation

A test network called *Uni Junø* network [18] with 30 validator nodes was used to evaluate the transaction time of the encrypted prescription items (after encryption step in Figure 1). The steps to automate the sending of transactions to the network were implemented in a shell script. All software and contracts developed for model evaluation are available on GitHub [16].

Transaction time for Smart Contracts in CosmWasm: Table I shows the transaction validation time for each prescription ranging from 0.92 kB to 130.50 kB containing all items

TABLE I
TRANSACTION VALIDATION TIME FOR THE CREATE_PRESCRIPTION CONTRACT METHOD USING COSMWASM IMPLEMENTATION WITH *Uni Junø* TESTNET

	Transaction time (in seconds)			
Number of Transactions	**Max.**	**Min.**	**Avg.**	**Std.**
300	6.26	1.50	2.69	0.71

(i.e., patient's personal information, medication, and diagnosis). Time refers to the Tendermint consensus process with inclusion in a block. On average, the time for a transaction to be validated by contract method took 2.69 seconds, with the maximum and minimum time being 6.26 seconds and 1.50 seconds, respectively. The variation in transaction time is due to the consensus delay, including peer-to-peer messaging between validator nodes.

VI. CONCLUSIONS

Real-world multi-stakeholder applications such as e-prescription and supply chain deal with digital and sensitive data, demanding privacy protection, consent management, data provenance, and transparency. We have presented a decentralized data governance framework for e-prescription that uses proxy re-encryption and smart contracts to let data owners control and manage their data through a trusted and transparent blockchain platform. We have shown how the data owners can record all the access requests and consents in an immutable ledger to monitor data lineage. Our proof-of-concept implementation uses CosmWasm and pyUmbral proxy re-encryption library to assess the feasibility and performance. Our evaluation results show that the proposed architecture can protect data owners' privacy and govern sensitive data access with minimal overhead. We believe that our data governance framework is beneficial to all multi-stakeholder applications that deal with sensitive and private digital data.

ACKNOWLEDGEMENT

Jo Ueyama would like to thank the Brazilian Research Council (CNPq) and São Paulo Research Foundation (FAPESP) for financing the bulk of his research, FAPESP grant ID 2013/07375-0. Rodrigo D. Garcia would like to acknowledge the financial support by the Brazilian Research Council (CNPq, process 133470/2020-2) for his research.

REFERENCES

[1] Bader Aldughayfiq and Srinivas Sampalli. Digital health in physicians' and pharmacists' office: A comparative study of e-prescription systems' architecture and digital security in eight countries. *OMICS: A Journal of Integrative Biology*, 25(2):102–122, 2021.

[2] Mahnaz Samadbeik, Maryam Ahmadi, Farahnaz Sadoughi, and Ali Garavand. A copmarative review of electronic prescription systems: Lessons learned from developed countries. *Journal of research in pharmacy practice*, 6(1):3, 2017.

[3] Shunrong Jiang, Haiqin Wu, and Liangmin Wang. Patients-controlled secure and privacy-preserving ehrs sharing scheme based on consortium blockchain. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2019. doi: 10.1109/GLOBECOM38437.2019.9013220.

[4] Phil Skolnick. The opioid epidemic: Crisis and solutions. *Annual Review of Pharmacology and Toxicology*, 58(1): 143–159, 2018. doi: 10.1146/annurev-pharmtox-010617-052534. URL https://doi.org/10.1146/annurev-pharmtox-010617-052534. PMID: 28968188.

[5] Rachel W. Faller, Jennifer Toller Erausquin, and Thomas P. McCoy. Misuse of prescription and illicit drugs in middle adulthood in the context of the opioid epidemic. *Substance Use & Misuse*, 56(2):333–337, 2021. doi: 10.1080/10826084.2020.1858107. URL https://doi.org/10.1080/10826084.2020.1858107. PMID: 33325317.

[6] Ehab Zaghloul, Tongtong Li, and Jian Ren. Security and privacy of electronic health records: Decentralized and hierarchical data sharing using smart contracts. In *2019 International Conference on Computing, Networking and Communications (ICNC)*, pages 375–379, 2019. doi: 10.1109/ICCNC.2019.8685552.

[7] May Alnafrani and Subrata Acharya. Securerx: A blockchain-based framework for an electronic prescription system with opioids tracking. *Health Policy and Technology*, 10(2):100510, 2021. ISSN 2211-8837. doi: https://doi.org/10.1016/j.hlpt.2021.100510. URL https://www.sciencedirect.com/science/article/pii/S2211883721000332.

[8] Rodrigo Dutra Garcia, Gabriel Augusto Zutião, Gowri Ramachandran, and Jo Ueyama. Towards a decentralized e-prescription system using smart contracts. In *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 556–561, 2021. doi: 10.1109/CBMS52027.2021.00037.

[9] Patrick Li, Scott D. Nelson, Bradley A. Malin, and You Chen. Dmms: A decentralized blockchain ledger for the management of medication histories. *Blockchain in Healthcare Today*, 2, Jan. 2019. doi: 10.30953/bhty.v2.38. URL https://blockchainhealthcaretoday.com/index.php/journal/article/view/38.

[10] DAVID Nunez. Umbral: a threshold proxy re-encryption scheme. *NuCypher Inc and NICS Lab, University of Malaga, Spain*, 2018.

[11] Renpeng Zou, Xixiang Lv, and Jingsong Zhao. Spchain: Blockchain-based medical data sharing and privacy-preserving ehealth system. *Information Processing & Management*, 58(4):102604, 2021.

[12] Kumar Bhaskaran, Peter Ilfrich, Dain Liffman, Christian Vecchiola, Praveen Jayachandran, Apurva Kumar, Fabian Lim, Karthik Nandakumar, Zhengquan Qin, Venkataraman Ramakrishna, Ernie GS Teo, and Chun Hui Suen. Double-Blind Consent-Driven Data Sharing on Blockchain. *2018 IEEE International Conference on Cloud Engineering (IC2E)*, pages 385–391, 2018. doi: 10.1109/ic2e.2018.00073.

[13] Matt Blaze, Gerrit Bleumer, and Martin Strauss. Divertible protocols and atomic proxy cryptography. In Kaisa Nyberg, editor, *Advances in Cryptology — EUROCRYPT'98*, pages 127–144, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg. ISBN 978-3-540-69795-4.

[14] Sijia Zhao and Donal O'Mahony. Bmcprotector: A blockchain and smart contract based application for music copyright protection. In *Proceedings of the 2018 International Conference on Blockchain Technology and Application*, pages 1–5, 2018.

[15] NuCypher. pyumbral pre. https://github.com/nucypher/pyUmbral, 2021.

[16] R.D Garcia. E-prescription model. https://github.com/rodrigodg1/e-prescription, 2021.

[17] English Prescribing Dataset. English prescribing dataset (epd). https://opendata.nhsbsa.net/dataset/english-prescribing-data-epd, 2021.

[18] CosmWasm. Cosmwasm testnets. https://github.com/CosmWasm/testnets, 2021.