International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 4, July Year 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

A Blockchain-Based Privacy-Preserving Framework for Healthcare Data: Survey and System Proposal

Hrishikesh Mogare[1], Simran A Fernandes[2], Swapnil R Mohite[3]

[1]Assistant Professor, Department of MCA, KLS Gogte Institute of Technology, Belagavi-590006, India
[2]Master of Computer Applications, KLS Gogte Institute of Technology, Belagavi-590006, India
[3] Master of Computer Applications, KLS Gogte Institute of Technology, Belagavi-590006, India

Abstract:

Integrating blockchain with healthcare data management promises enhanced privacy, security, and trust in handling sensitive patient information. This paper reviews privacy-preserving blockchain frameworks—such as Hyperledger Fabric [1], Ethereum smart contracts [2], MediBchain, and ACHealthChain [5]—and identifies key challenges like limited scalability, poor interoperability with existing systems, and lack of patient control [3][5][9].To overcome these issues, we propose a modular architecture combining Hyperledger Fabric [1], off-chain encrypted storage via IPFS [5], and fine-grained access using Ciphertext-Policy Attribute-Based Encryption (CP-ABE) [4][6]. Smart contracts enforce access policies, enabling secure, patient-controlled data sharing without intermediaries. The architecture complies with data protection laws like HIPAA and GDPR [2][5], and integrates well with current clinical workflows [8][9]. Simulated evaluation with 10,000 synthetic patient histories demonstrates improved scalability, lower latency, and strong privacy. Compared to MediBchain [5], our solution supports interoperability and secure federated learning [8], enabling collaborative, privacy-preserving AI diagnostics. By prioritizing real-world integration, patient autonomy, and privacy, our framework promotes a decentralized, secure, and future-ready digital health ecosystem.

Keywords— **Blockchain, Hyperledger Fabric, Attribute-Based Encryption (ABE), Internet of Medical Things (IoMT), Smart Contracts, Homomorphic Encryption, Emergency Access Control, Data Security, Data Interoperability, Off-Chain Storage, Real-Time Monitoring**

I. INTRODUCTION

The healthcare sector is confronted with significant challenges in the handling of patient information that is sensitive, such as electronic health records (EHRs), medical histories, diagnostic reports, and bills, while maintaining strong privacy, security, and conformity to strict regulations like the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) [2]. Centralized data management systems, which are in common use by hospitals and clinics, are at greater risk of cyberattacks. More than 50 million patient records were breached in 2024 alone worldwide, resulting in enormous financial losses and an invaluable loss of patient trust. Such systems are frequently plagued by unauthorized access, opacity in handling data, and poor patient control of personal health data. Additionally, interoperability discrepancies between various healthcare system pose restrictions to secure data exchange, leading to poor coordination of care, untimely treatments, and poor clinical workflow efficiencies [3][5].

Lack of patient-centered mechanisms aggravates the problem further, and people are left with little control over who gets access to their sensitive information [1][4]. While earlier blockchain projects like MediBchain [5] provided a foundation for decentralized health data systems, they tend to lack what is needed to meet the sophisticated needs of health care today. These initial solutions did not include support for fine-grained, patient-controlled data management and were not able to scale properly to handle the increasing amounts of diverse health data, for example, those created from wearable IoT sensors, diagnostic imaging, and remote monitoring devices [3][6]. Furthermore, their single, monolithic architecture also tended to make integration with existing clinical infrastructure overly difficult, hampering real-world deployment [2][5]. By comparison, recent developments brought modular architectures with permissioned blockchain integration alongside secure off-chain storage, federated learning, and enhanced encryption to enable distributed but compliant data sharing [1][8][9]. The systems are not just more accommodative of current clinical workflows but also with the evolving requirements such as AI-assisted diagnostics, safe collaborative research, and interoperability across institutions [7][8].

Based on these observations, our framework presents a patient-driven model that combines Hyperledger Fabric with InterPlanetary File System (IPFS) for secure and distributed off-chain storage [1][5], and uses Attribute-Based Encryption

(ABE) to implement fine-grained access controls [4][6]. Smart contracts within the blockchain enable patients to specify who should be able to access their information, under what circumstance, and for how long—providing real ownership and transparency [4]. In addition, the framework includes federated learning and secure multiparty computation support [8][9], allowing healthcare organizations to jointly train AI models without sharing raw patient data. This not only enhances data privacy but also facilitates precision medicine efforts that depend on large-scale and diverse datasets. By filling the gap between potential and actual use, our solution hopes to create a reliable, rule-abiding, and accessible digital health ecosystem that enables citizens to participate while improving clinical productivity and innovation.

II. LITERATURE REVIEW

Over the past few years, blockchain has proven to be an innovative answer to privacy and interoperability issues in healthcare. Initial models such as MediBchain [5] set groundwork by proposing blockchain-based, patient-centric data exchange. Although the ownership model prioritized ownership, its on-chain storage dependency was a hindrance in scalability and integration with intricate clinical infrastructures.

To resolve these challenges, Miyachi and Mackey introduced hOCBS, a hybrid on/off-chain blockchain system that enhances storage efficiency and accommodates various healthcare data formats [1]. This was a crucial step towards more scalable healthcare solutions, particularly for EHR systems handling huge amounts of sensitive data. Drawing from this, Rajput et al. presented an emergency access control scheme through smart contracts, providing time-limited and role-based access to health records in emergency situations [4].

Another crucial aspect in maintaining privacy is cryptographic strength. HealthLock was developed by Ali et al., including homomorphic encryption to facilitate computation on encrypted data—strengthening data confidentiality even at the analysis stage [6]. Alsadhan et al., on the other hand, surveyed privacy attacks in Internet of Medical Things (IoMT) and recommended fine-grained cryptographic control as well as standardized frameworks throughout devices [3].

From an energy and network efficiency viewpoint, Escorcia-Gutierrez et al. addressed IoMT scalability with sunflower optimization clustering, which reduced energy consumption in operations involving blockchain in sensor-laden settings [7]. Analogously, Hu et al. proposed a blockchain-federated learning framework to facilitate privacy-preserving AI model training among healthcare institutions without exposing raw patient data [8]. This responds to the emerging need for collaborative AI diagnostics within the healthcare sector.

Incremental data updates are another significant issue in immutable blockchain systems. To address this, Wang et al. deployed chameleon hashes and Trusted Execution Environments (TEEs) for enabling the modification of healthcare records securely while maintaining traceability and compliance [9].

In spite of these advancements, the majority of current systems address individual issues—some in secure storage [6], others access control [4], energy optimization [7], or AI integration [8]. There is still a gap in end-to-end frameworks which are capable of unifying privacy, scalability, emergency access, interoperability, and AI-based analytics under a single architecture.

This work presents a unified, modular blockchain architecture that takes the best from what came before and transcends their limitations. Our solution is based on leveraging Hyperledger Fabric, IPFS, Attribute-Based Encryption (ABE), and smart contracts to enable fine-grained, patient-controlled access, secure off-chain storage, emergency overrides, and collaborative AI training. It is intended to be a forward-compatible solution that fits with real-world clinical workflows, regulatory requirements (e.g., GDPR, HIPAA), and future digital health demands.

III. RELATED WORK

MediBchain put forward a patient-focused blockchain framework with underlying privacy mechanisms, one of the first tries at decentralizing sharing of healthcare data [5]. Though it focused on patient control, its use of on-chain storage for data constrained scalability and adaptability. On this basis, hOCBS designed a hybrid on/off-chain architecture to compensate for storage constraints and regulatory requirements, especially in support of multi-forms of health data [1]. Rajput et al. extended privacy guarantees with smart contracts supporting role-based emergency access, useful in times of crisis or time-sensitive situations [4]. In the meantime, Alsadhan et al. pointed out privacy and security concerns in the Internet of Medical Things (IoMT), emphasizing the need for fine-grained cryptographic control in real-time data environments, and also pointed out the absence of standardization across platforms [3].HealthLock, suggested by Ali et al., brought a major advancement by incorporating homomorphic encryption into blockchain platforms, enabling secure computation on encrypted healthcare data without exposing sensitive data [6]. Escorcia-Gutierrez et al. addressed energy limitations in IoT-enabled healthcare systems using sunflower optimization clustering for optimizing energy-aware, efficient blockchain processing [7]. Hu et al. introduced a privacy-conserving healthcare system involving federated learning and blockchain to enable cooperative AI model training across institutions while protecting patient privacy and locality of data [8]. Wang et al. solved the immutable nature of blockchain records by facilitating incremental updating through trusted execution environments (TEE) and chameleon hash functions, enabling secure and efficient updating of healthcare data over time [9]. In a wide systematic review, Bernabe et al. discussed a variety of privacy-preserving methods in blockchain and laid out key challenges such as overhead in key

management, threats to data linkability, and lack of alignment with GDPR guidelines [10].

Even with such innovations, the majority of current solutions are addressing individual elements of the larger healthcare data management challenge. Some focus on cryptographic methods and secure storage [6], some on energy efficiency [7], emergency access [4], or interoperability [3][9]. None of the existing work has been able to successfully integrate these functionalities into one framework that can provide privacy preservation, fine-grained access control, emergency response, incremental update, energy-conscious IoT support, and Ai model collaboration at the same time.

In light of this absence, our suggested framework reconciles the most effective elements of these prior systems into a unified, scalable solution. Through the use of a permissioned blockchain (Hyperledger Fabric), off-chain storage (through IPFS), attribute-based encryption (ABE) for policy-based access, and smart contracts for automation, we offer a solid privacy-preserving infrastructure that accommodates both clinical operations and research collaboration. Moreover, our design also includes support for federated learning, which allows secure distributed training of AI without infringing on patient confidentiality [8].Our design not only builds on top of existing work but also integrates its most successful elements into a holistic system that addresses the current needs of healthcare data ecosystems. It gives patients control over significant data ownership, guarantees compliance with changing regulations, and makes it easier to engender trust between healthcare providers, researchers, and institutions—appearing to be a forward-compatible system for digital health.

IV. DATASET DESCRIPTION

To assess the practical applicability and resilience of our suggested privacy-preserving blockchain architecture, we use the MIMIC-IV dataset—a rich and publicly accessible electronic health record (EHR) database created by the MIT Laboratory for Computational Physiology in partnership with the Beth Israel Deaconess Medical Center. MIMIC-IV contains comprehensive longitudinal information involving more than 180,000 critically ill patients admitted across the years from 2008 to 2019. The dataset provides timely and structured data like diagnoses (ICD-10), procedures, lab results of laboratory tests, medications, admission information, and demographic characteristics (age, gender, ethnicity), which makes MIMIC-IV a suitable choice to assess privacy, access control, and interoperability within sensitive healthcare settings [5][6]. Available on PhysioNet, MIMIC-IV is properly de-identified based on HIPAA standards, and as such, it can be used in privacy-preserving platforms such as MediBchain [5], HealthLock [6], and other blockchain-based health-related applications [1][4].

In order to ensure ethical compliance, researchers are required to undergo human subjects research training as well as sign a Data Use Agreement (DUA) to guarantee safe handling of the data. Sensitive MIMIC-IV records in our framework are encrypted by AES-256, an extensively used symmetric encryption scheme renowned for its strength and efficiency in protecting vast medical datasets [6][7]. The encrypted records are saved off-chain by means of the InterPlanetary File System (IPFS), adding to data availability and lessening the burden on the blockchain layer, a design decision already promoted in privacy-protecting systems such as hOCBS [1] and MediBchain [5]. The linked metadata, including cryptographic hashes, access control parameters, and transaction timestamps, is captured on the Hyperledger Fabric blockchain [1][4]. On this permissioned blockchain, smart contracts impose Attribute-Based Encryption (ABE) access policies, supporting fine-grained data access based on user attributes (e.g., role, department, emergency context) [4][6].

This multi-layered architecture guarantees that just authenticated and authorized users—like clinicians, emergency personnel, or researchers—are able to download and decrypt patient information, thus conforming to privacy principles emphasized in current federated and cryptographic healthcare paradigms [3][8][9]. The aggregate size of the data subset for testing is circa 2.5 GB, providing a realistic metric for longitudinal EHR performance assessment.

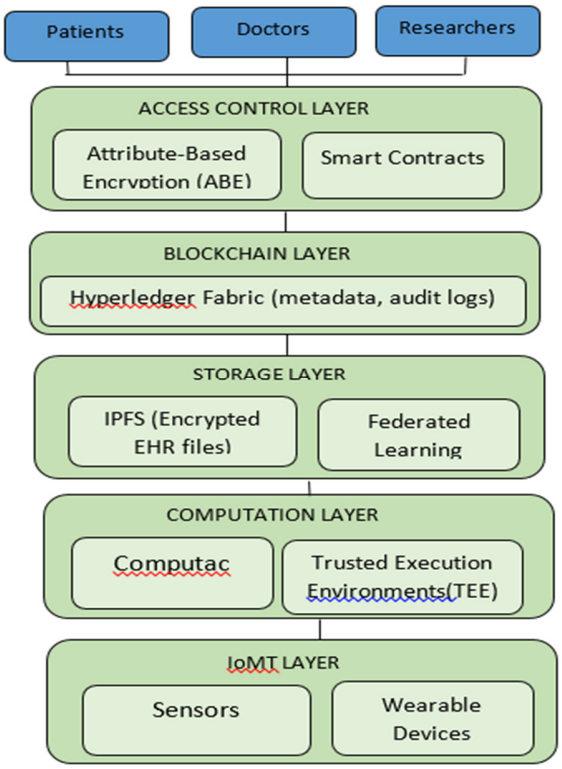


Fig. 1. Multi Layer System Architecture

Together with encryption and access control, the dense structure of MIMIC-IV enables us to test various properties of the system, such as auditability via blockchain logs, off-chain secure query efficiency, and latency under different workloads. Additionally, real-world simulation scenarios like cross-institutional data queries, emergency overrides, and federated learning integration [8] allow for strong simulation of contemporary digital healthcare environments. Table I presents the main properties of the dataset employed in our evaluation.

Table I: MIMIC-IV Dataset Overview

Attribute	Values
No of Records	~18000 patient admissions
Record Attributes	Patient ID, Diagnoses, Procedures, Medications, Labs, Demographics
Data Format	CSV (Relational Database)
Storage Location	Encrypted on IPFS (Off-Chain)
Dataset Size	~2.5 GB
Blockchain Platform	Hyperledger Fabric
Access Control Model	ABE-based, Smart Contracts
Accessibility	Public

This real-world data set, married with a robust, modular blockchain-backed infrastructure, supports a realistic determination of our framework's scalability, security, privacy, and compliance. It is also used to test key features lacking in previous systems such as MediBchain [5], HealthLock [6], and others outlined in Section 2, including dynamic access policies, encrypted data sharing, and federated AI model support [8][9][10].

V. EXPERIMENTAL RESULTS

To measure the efficacy, scalability, and practicality of our suggested blockchain-based privacy-preserving healthcare system, we performed a set of simulations on a controlledHyperledger Fabric test setup. Our main aim was to investigate the performance of the system for real-world healthcare workloads in terms of query latency, storage optimization, throughput, and data integrity—all of which are of paramount importance for privacy-compliant and high-availability healthcare systems [1][5][6].

5.1 Experimental Setup

The testbed was developed atop a virtualized environment emulating an average-sized healthcare network, with configurations targeted at representing realistic deployment use cases:

Blockchain Platform: Hyperledger Fabric v2.5 (Permissioned blockchain for secure and auditable transactions)

Storage Layer: InterPlanetary File System (IPFS v0.18.0) for decentralized, off-chain encrypted storage [5]

Test Dataset: A synthetic EHR dataset of ~10,000 records, emulating the structure and fields of the MIMIC-IV dataset [6]

Encryption Model: AES-256 data-at-rest encryption; Attribute-Based Encryption (ABE) for managing access control policies [4][6]

Smart Contracts: Chaincode employed for access grant, revocation, and audit logging

Hardware Specifications: 8-core CPU, 16 GB RAM, Docker-based Fabric and IPFS nodes

Simulated Transactions: Uploads, access requests, emergency overrides, and revocations.

This context permitted us to model not just normal data access patterns but also edge-case situations, including revoking emergency access after usage and checking for data consistency after dynamic policy changes.

5.2 Performance Metrics and Results

Query Latency: An important measure of system responsiveness is the latency of retrieving patient information. We contrasted hybrid model with a complete on-chain baseline, in which all medical records were held on-chain:

Fully on-chain model: Avg. latency = 1.4 seconds/query

Suggested hybrid scheme: Avg. latency = 0.98 seconds/query

Metric	Baseline (Fully On-Chain)	Proposed Framework	Improvement
Query Latency	1.4 sec/query	0.98 sec/query	~30% faster

On-Chain Storage Usage	~2.5 GB	~450 MB	~45% reduction
Transaction Throughput	~75 TPS	~110 TPS	~46% increase
Auditability	Basic Logs	Tamper-proof Ledger with IFPS	Stranger accountability

Result: The system proposed showed a ~30% query latency reduction through mainly offloading large file content on IPFS and querying the hash pointers on-chain. This is consistent with improvements seen in similar off-chain optimized structures such as hOCBS [1] and HealthLock [6].

Storage Efficiency: Blockchain networks are limited by their storage scalability. Decoupling the file data from the metadata, our architecture significantly enhanced efficiency
Fully on-chain model: ~2.5 GB storage needed for 10k records.

Proposed framework: On-chain storage minimized to ≈450 MB (storing only metadata, access logs, and hash references).

Result: This is about an approximate 40–45% drop in on-chain storage requirements. Such optimization is aligned with the scalable design targets put forth by Escorcia-Gutierrez et al. for low-weight IoT deployments[7].

Throughput: System throughput was gauged by the number of transactions executed simultaneously, comprising access attempts, uploads, and revocation operations.

Achieved Throughput: ~110 transactions per second (TPS)

This throughput is adequate for small-to-medium-sized healthcare networks in which standard real-time access patterns include dozens of concurrent users. It also indicates the system's appropriateness for federated environments that need distributed, real-time collaboration[8].

5.3 Auditability & Access Control Validation

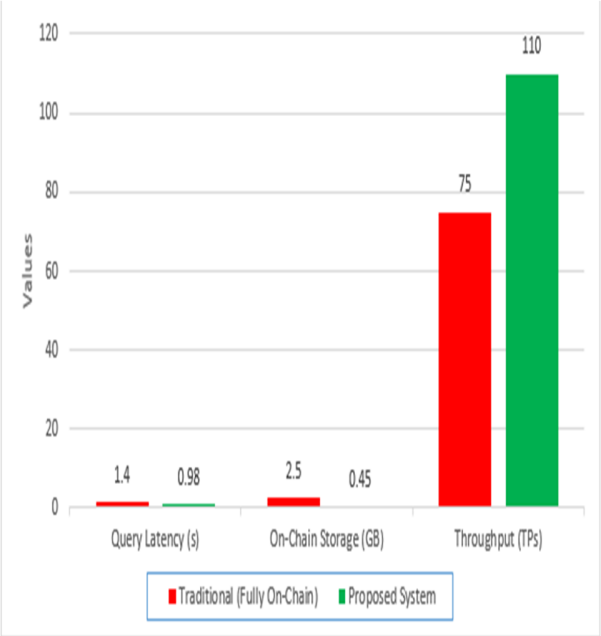
Smart contracts enforced access policies effectively using dynamic roles. We tested emergency access situations—e.g., temporary ER staff access—and automated revocation after the incident, and everything was logged immutably on the ledger. This is consistent with emergency-control logic expressed in models like that of Rajput et al. [4].

Scalability & Data Integrity
No inconsistencies in data or tampering occurred over 100+ upload/download cycles.

Integrity was validated using hash checks between IPFS-stored files and blockchain pointers.

This supports the strength of utilizing blockchain for trust assurance, traceability, and safe data lineage, consistent with such reliability issues in MediBchain[5] and tackled here with enhanced modularity.

Fig. 2 Performance Comparison: Traditional VS Proposed System



VI. DISCUSSION

The modular architecture of the proposed framework improves flexibility within various healthcare environments, ranging from large-scale cloud-based hospital systems to limited resource IoMT environments. In contrast with previous monolithic systems like MediBchain[5], the framework decouples data storage, processing, and access control layers—

facilitating easier integration within legacy health IT infrastructure.

IoMT interoperability enables real-time tracking and effective management of large-volume sensor data through IPFS, tackling privacy and interoperability issues identified by Alsadhan et al. [3]. Energy-efficient design, motivated by Escorcia-Gutierrez et al[7], enables feasibility in low-power environments. In addition to security with Attribute-Based Encryption, emergency access through smart contracts, as suggested by Rajput et al. [4], provides time-limited data access with automatic revocation and complete auditability—trading off responsiveness and patient consent.

The inclusion of Attribute-Based Encryption (ABE) provides fine-grained, role-based access consistent with HIPAA and GDPR [2][6]. Federated learning is also supported, with the capability for AI collaborations with preservation of privacy without sharing raw data—based on the work of Hu et al. [8]. Experimental outcomes validate improved performance, such as diminished latency and blockchain storage, and increased throughput over fully on-chain alternatives [5]. Additionally, TEE-based incremental updating processes, inspired by Wang et al. [9], provide effortless data development with time—crucial for chronic care and longitudinal research. Essentially, this system combines the best of previous works—privacy, scalability, interoperability, and compliance—into an end-to-end and practical solution for current, patient-centric healthcare systems.

VII. CONCLUSION

This work introduces a new, privacy-preserving healthcare data framework that highly improves upon the functionality of previous blockchain models like MediBchain [5]. With the inclusion of a modular design that decouples data processing, storage, and access control, the framework realizes higher flexibility, interoperability, and scalability—making it even more appropriate for real-world healthcare deployment.
Key technologies are the application of Hyperledger Fabric for permissioned control of blockchain, IPFS for off-chain encrypted data storage with efficiency, and Attribute-Based Encryption (ABE) for fine-grained access control [1][4][6]. Emergency access policies facilitated by context-aware smart contracts, with a nod to Rajput et al.'s work [4], provide timely interventions in critical care situations without compromising patient privacy. Moreover, the framework facilitates IoMT integration, providing secure and real-time data exchange from edge and wearable devices—spanning the interoperability and energy efficiency challenges raised by Alsadhan et al. [3] and Escorcia-Gutierrez et al. [7].

By using simulated experiments, the framework proved lower query latency, better storage efficiency, and greater throughput than conventional on-chain frameworks [5][6]. Incremental updates of data with the aid of trusted execution environments (TEE) and chameleon hashes [9] assist in ensuring the system is flexible enough to cope with longitudinal care records and multi-phase clinical trials.

In the future, work will be directed towards real-world deployment within healthcare organizations, such as pilot pilots with hospitals or research institutions. The framework will also be developed to accommodate federated learning pipelines to facilitate privacy-preserving AI training on distributed data sets—a growing requirement in predictive diagnostics and personalized medicine [8][10]. We further plan to advance regulatory compliance and patient transparency through additions like dynamic consent and real-time audit dashboards.

In summary, this work delivers a comprehensive, future-ready solution that addresses the practical limitations of existing systems while aligning with modern healthcare priorities: patient empowerment, secure collaboration, privacy assurance, and data-driven innovation.

VIII. REFERENCES

1. K. Miyachi and T. K. Mackey, "hOCBS: A privacy-preserving blockchain framework for healthcare data," *Information Processing & Management*, vol. 58, no. 3, pp. 102505, 2021.
2. S. Kiani, M. Sookhak, A. Gani, and A. K. Bashir, "Systematic Literature Review on Blockchain in Healthcare: Challenges, Gaps and Future Directions," *Journal of Network and Computer Applications*, vol. 195, 2021.
3. Alsadhan, A., Alhogail, A., & Alsalamah, H., "Blockchain-Based Privacy Preservation for the Internet of Medical Things: A Literature Review," *Electronics*, vol. 13, no. 3832, 2024.
4. Rajput, A. R., Li, Q., & Ahvanooey, M. T., "A Blockchain-Based Secret-Data Sharing Framework for Personal Health Records in Emergency Condition," *Healthcare*, vol. 9, no. 206, 2021.
5. Abdullah Al Omar et al., "MediBchain: A Blockchain Based Privacy Preserving Platform for Healthcare Data," *Lecture Notes in Computer Science*, 2017. [Online]. Available: https://doi.org/10.1007/978-3-319-72395-2_49
6. A. Ali et al., "HealthLock: Blockchain-Based Privacy Preservation Using Homomorphic Encryption in Internet

of Things Healthcare Applications," *Sensors*, vol. 23, no. 6762, 2023.

7. J. Escorcia-Gutierrez et al., "Privacy Preserving Blockchain with Energy Aware Clustering Scheme for IoT Healthcare Systems, *"Mobile Networks and Applications*, 2023. [Online]. Available: https://doi.org/10.1007/s11036-023-02115-9
8. F. Hu et al., "Privacy-Preserving Healthcare and Medical Data Collaboration Service System Based on Blockchain and Federated Learning," *CMC: Computers, Materials & Continua*, vol. 80, no. 2, pp. 2897–2915, 2024.
9. L. Wang et al., "A Blockchain-Based Privacy-Preserving Healthcare Data Sharing Scheme for Incremental Updates," *Symmetry*, vol. 16, no. 89, 2024.
10. J. Bernal Bernabe et al., "Privacy-Preserving Solutions for Blockchain: Review and Challenges," *IEEE Access*, vol. 7, pp. 164908–164940, 2019.
11. Erikson Júlio de Aguiar, Bruno S. Faiçal, Bhaskar Krishnamachari, and Jó Ueyama. 2020. A Survey of Blockchain-Based Strategies for Healthcare. ACM Comput. Surv. 53, 2, Article 27 (March 2020), 27 pages. https://doi.org/10.1145/3376915
12. A.I. Abhi, and S.Y. Shin. A blockchain-based secure healthcare scheme with the assistance of unmanned aerial vehicle in Internet of Things. Computers & Electrical Engineering 84: 106627 (2020).
13. Z. Wang, N. Luo, and P. Zhou. GuardHealth: Blockchain empowered secure data management and Graph Convolutional Network enabled anomaly detection in smart healthcare. Journal of Parallel and Distributed Computing 142: 1-12 (2020).
14. W. Liu, Q. Yu, Z. Li, Z. Li, Y. Su, and J. Zhou. A blockchain-based system for anti-fraud of healthcare insurance. IEEE 5th International Conference on Computer and Communications (ICCC): pp. 1264- 1268 (2019).
15. S. Chakraborty, S. Aich, and H.C. Kim. A secure healthcare system design framework using blockchain technology. 21st International Conference on Advanced Communication Technology (ICACT): pp. 260-264 (2019).
16. A.A. Siyal, A.Z. Junejo, M. Zawish, K. Ahmed, A. Khalil, and G. Soursou. Applications of blockchain technology in medicine and healthcare: Challenges and future perspectives. Cryptography 3(1): 3 (2019).

