A Combined Framework of InterPlanetary File System and Blockchain to Securely Manage Electronic Medical Records

Abdullah Al Mamun, Md. Umor Faruk Jahangir, Sami Azam, M. Shamim Kaiser, and Asif Karim

Abstract Blockchain has become a popular research area since its introduction, as the benefit has been seen in various industries. It could greatly benefit the healthcare sector as it offers anonymity, immutability, and decentralization of data. Electronic Medical Record (EMR) systems face crucial problems regarding data security, accessibility, and management. A great deal of security threats relating to patient privacy involves unauthorized access to medical records, misuse of patient's disease reports, and so on. To address these issues, we have proposed a blockchain combined with the InterPlanetary File System solution framework for EMR in the healthcare industry. The aim is to implement the blockchain for EMR and provide access rules for various users of it. The proposed framework, while protecting patient privacy, allows convenient access by approved authorities such as healthcare providers to medical data.

Keywords Blockchain · IPFS · EMR · Data Security · SHA256 · Python · Postman · Flask

A. Al Mamun (✉) · M. S. Kaiser
Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh
e-mail: abdullah.iiuceee@gmail.com

M. S. Kaiser
e-mail: mskaiser@juniv.edu

Md. U. Faruk Jahangir
Department of Computer Science & Engineering, Shahjalal University of Science and Technology, Sylhet, Bangladesh
e-mail: jahangircs.sust@gmail.com

S. Azam · A. Karim
College of Engineering, IT and Environment, Charles Darwin University, Casuarina, Australia
e-mail: sami.azam@cdu.edu.au

A. Karim
e-mail: asif.karim@cdu.edu.au

M. S. Kaiser et al. (eds.), *Proceedings of International Conference on Trends in Computational and Cognitive Engineering*, Advances in Intelligent Systems and Computing 1309, https://doi.org/10.1007/978-981-33-4673-4_40

1 Introduction

Technology has a significant role to play in addressing diverse issues in our everyday lives. The necessity of technological development in the medical sector is the demand of time now. The healthcare sectors produce a high volume of medical data every day that need to be stored, disseminated, and accessed daily. For example, medical data is created when a patient goes through some medical tests, such as MRI, X-Ray, and USG, or when a doctor prescribes the patient. However, healthcare data are breached every day in different parts of the world. For instance, the U.S. Department of Health and Human Services (HHS) stated in one of their reports in 2018 that 6.1 million people are victims of 229 healthcare data breaches, and the report was published in the civil rights of breach portal. In 2014, Statistica mentioned about 783 data breaches, with around 85.61 million records were compromised, which was approximately 500% since 2005 [20, 34].

Healthcare industries have management complexity and diversity of providers. Point to be noted that the medical record of a patient is stored in a hospital in electronic form, and the patient doesn't get accessed to it when taking treatment from another hospital [21, 24, 29]. People are dying due to the lack of proper infrastructure, human resources, funding, and data preservation policies [28]. Moreover, Electronic Medical Records (EMR) are important and extremely confidential private information for healthcare treatment like HIV and cancer. These data need to be shared among various peers, including healthcare providers, insurance companies, researchers, passport officials, law enforcement agencies, and relatives. Maintaining security and stacking up previous diagnosis reports of these valuable data is crucial [4, 8].

For consultation and treatment, a patient might visit multiple medical institutions or move from one hospital to another. As a result, a customer does not get all his medical data accumulated when needed at various locations. These medical records are generally too sensitive to fall into the wrong hand. Usually, these data are stored in various organizations' private databases, so these data remained centralized. Many researchers around the globe proposed several frameworks to address these issues. However, the recent solutions lack in reducing the security threats in healthcare data management [22, 30]. Hence, we proposed a framework based on the blockchain and IPFS network to tackle the problem.

The remainder of the paper is worded as follows : Section 2 outlined the relevant work on health data management and limitations. Section 3 addressed the background technology used in our proposed system. Section 4 outlined and applied the device model. Section 5 ultimately ended the paper and discussed the future continuation of the analysis.

2 Related Works

There are potential applications based on blockchain technology for healthcare data managing proposed by many researchers. Yue et al. claim to be the first to incorporate blockchain into healthcare system architecture [35]. They proposed the architecture of a gateway application for healthcare data for simple and safe monitoring and sharing of medical data between various entities. However, their proposed system has not been tested nor implemented. They only showed some possibilities of sharing data for study purposes without any security or privacy measures. Jenkins et al. suggested to use blockchain combining with supervised machine learning for medical large data analysis with some functional markers, which involves bio-metric and biomedical data [19].

MedRec [5] demonstrated how to apply decentralization principles in an EMR framework to large-scale data management. The authors attempted to illustrate an approach to accessing medical information via a detailed usage log. However, the prototype was not commercially applied. The scheme was based on the permissionless public blockchain. Dubovitskaya et al. explored possible applications of blockchain technologies in different healthcare environments [13]. The authors attempted to clarify the theoretical blockchain infrastructure to better handle healthcare records. They couldn't demonstrate any realistic application of any suggested framework. Omar et al. suggested a blockchain-based encryption network protecting patient records. The data is encrypted and processed in a federation blockchain, and to access the data [1], a data user must receive a data owner decryption key. A prototype of Hyperledger Medical data exchange between patient-doctor and pharmacist was seen in [32]. Using the AES256 encryption method, a patient encrypts the medical record, exchanging the keys with qualified users (doctors, pharmacists). However, as stated earlier, the proposed scheme was a project and had no functional implementation.

Shahnaz et al. suggested an ethereum-based blockchain algorithm to better handle EMR [31]. They wrote in Solidity the access management rules for healthcare records, measured costs, and evaluated performance for a theoretical case scenario. The authors have addressed EVM's average execution time, latency, and contract passage. They couldn't demonstrate any realistic case scenario or execution, though. The problems of treating patient medical information using blockchain are addressed of [28]. ABE and searchable ciphertext encryption technology are critical data protection, and fine-grained access control concerns technologies. A decentralized data management solution can solve the single failure point of traditional cloud storage systems. Combining Etherum and IPFS technologies may be a likely solution to storage problems [33]. People now switch to blockchain-based cloud data storage, as traditional cloud computing lacks encryption and fine-grained access control [14].

Recently, Kumar et al. suggested distributed off-chain storage of healthcare patient diagnosis reports in conjunction with IPFS. Authors suggested an idea to store medical records in the cloud with medical workers, and after some mining process, the block containing medical records would be created [23]. They have not considered, though, that anyone with IPFS hash still has access to medical records. Besides, they demonstrated no real-time implementation of the proposed method [25].

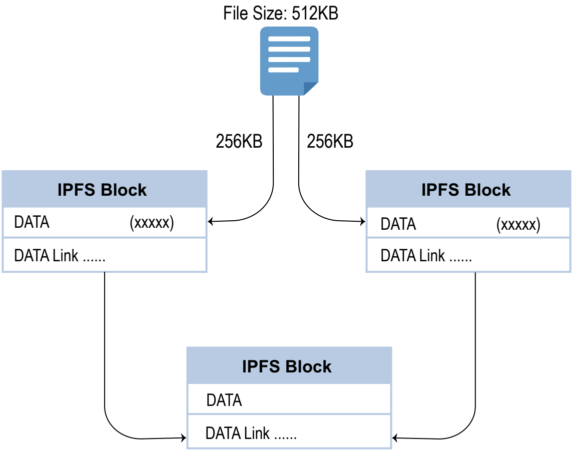
3 Background Technologies

3.1 IPFS

IPFS offers a distributed peer-to-peer storage structure to store and access encrypted massive volumes of EMR when required. When deleting redundant files from version-control history, IPFS stores files with their content-addressed hash [6] from a distributed hash table. The data contained in IPFS can only be modified by the hash value. Instead of keeping complete blockchain medical data, the IPFS uploads the hash value of the data to be transmitted. This would reduce mining latency and expense [2, 10, 12]. This stable IPFS property makes it ideal for storing vital and confidential clinical data. The cryptographic hash produced could be stored to minimize exhaustive blockchain computational operations. The IPFS protocol operates as: a unique content address is used to store and retrieve data on IPFS; The IPFS network do not allow the presence of duplicate files; and content exploration occurs using the DHTs [7].

IPFS is special in saving or accessing files over the traditional network. Although standard web uses location-based data processing addressing, IPFS uses content-based addressing. The key issue with conventional file storage or recovery is that it becomes unavailable to users if a file is deleted from the server location. However, IPFS overcome this problem by using the content address of data for accessing it. When a user saves a file in the IPFS network, IPFS splits files into several blobs, and an empty blob connects all [15] spluttered blocks (Fig. 1). Another great feature of the IPFS network is that it supports Git's document version control framework. A consumer can accurately monitor data changes over time.

Fig. 1 Data storing process of IPFS



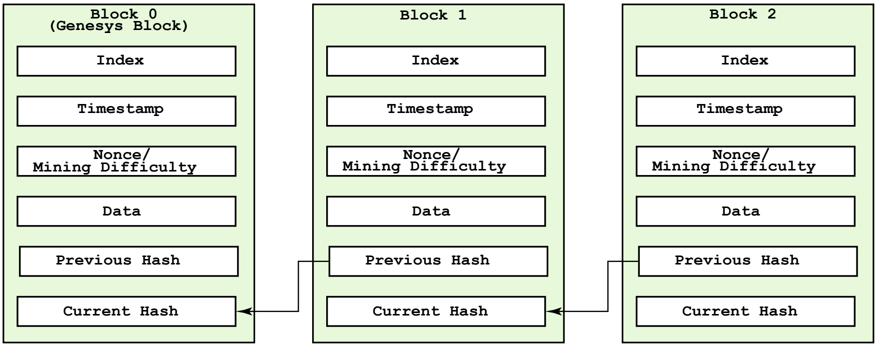


Fig. 2 Architecture of basic blockchain operation

3.2 The Blockchain

An individual (or group of people) known as Satoshi Nakamoto conceptualized the first blockchain in a digital currency form i.e. bitcoin in 2008. The primary aim was to solve the digital currency-related double-spending problem, but soon the technology began to use in several other [26] applications.

It uses a decentralized approach to disseminate and share information/data. The blockchain contains hyped, secured and peer-to-peer exchange packets. A blockchain has many advantages, including confidentiality, transparency, and data integrity, without third-party interference. These advantages make maintaining a patient's medical records equal, and technical advancement in the healthcare industry has prioritized patient health data protection. Several researchers find a viable solution [9, 17, 18] to use blockchain technologies in healthcare. Blockchain's simple architecture showed in Fig. 2 Several attributes exist, including block index, timestamp, nonce/mining complexity, details, current hash, previous hash, and more, in the blockchain.

4 System Model and Implementation

4.1 Proposed Workflow

The proposed workflow for EMR preservation and distribution has shown in Fig. 3. At first, a patient performs some medical diagnoses or tests at a healthcare center (Healthcare A) (1). Then, he receives the medical record or records from a healthcare authority (2). In step 3, the author encrypts the EMR using AES-256 algorithm [11, 16]. After that, the patient receives the IPFS hash by uploading the EMR to the IPFS network using our proposed system (4). He then uploads the hash of EMR with name in the blockchain network (5). The blockchain stores all the medical records

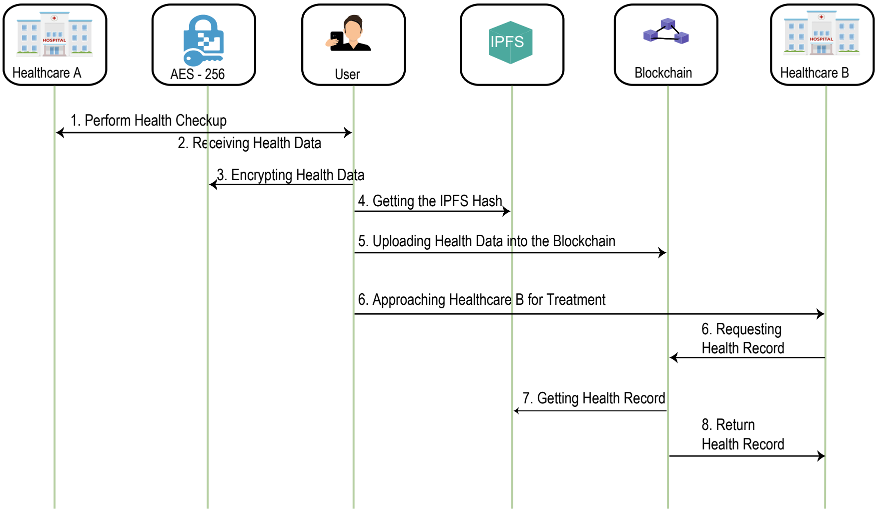

Healthcare A
AES - 256
User
IPFS
Blockchain
Healthcare B
1. Perform Health Checkup
2. Receiving Health Data
3. Encrypting Health Data
4. Getting the IPFS Hash
5. Uploading Health Data into the Blockchain
6. Approaching Healthcare B for Treatment
6. Requesting Health Record
7. Getting Health Record
8. Return Health Record

Fig. 3 Proposed workflow for secure storage and distribution of of EMR on blockchain

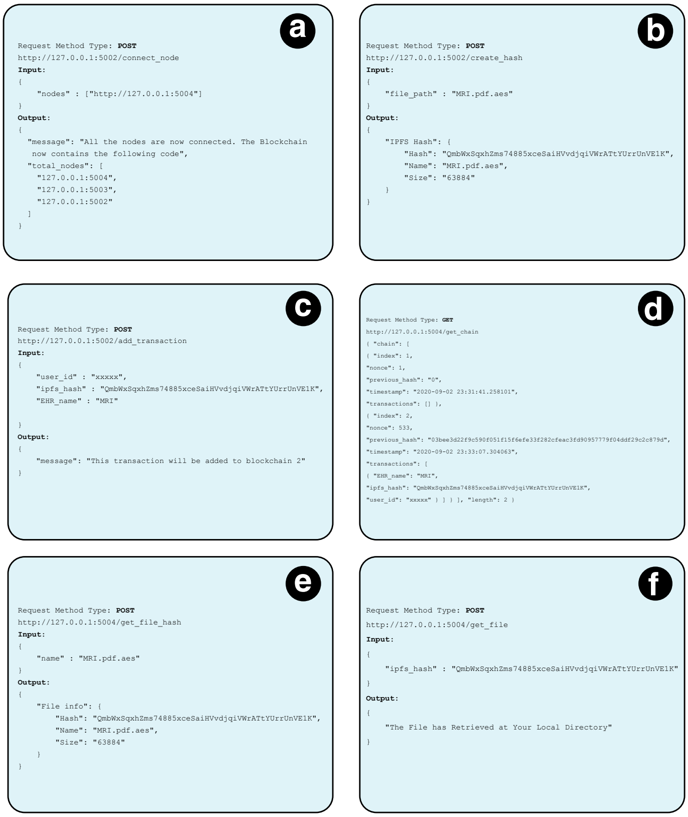
in different blocks. In the near future, when the patient needs to get treatment or consultation based on his recent symptoms and previous medical history, he can easily share the EMR from the blockchain with other healthcare authorities like pharmacies, doctors, or immigration agencies. At this stage, the authority will request for EMR on the blockchain using the user id and name of certificate (6). The blockchain network then retrieves the specific EMR from the IPFS network using the unique hash key for that records (7). Finally, the legitimate authority, for instance, healthcare B, will get a copy of EMR by decrypt it with the help of patient and take the necessary decision based on it (8).

4.2 Implementation

Experimental Setup: We used the Windows 10 operating system (version 1909) with 6 GB of RAM and 1 TB of HDD in our proposed model. We used Python (version 3.7.4) as a high-level programming language and coded it on Spyder IDE. We also used Flask (version 0.12.2) as the web application framework for our work. We deployed codes in the Postman API framework (version 7.29.1).

Connecting Peers: We had our work checked in two separate nodes Fig. 4a. Flask usually runs on http:/127.0.0.1:5000/. However, we used two separate Flask ports for our decentralization tests which are:

- http:/127.0.0.1:5002/
- http:/127.0.0.1:5003/



```
Request Method Type: POST
http://127.0.0.1:5002/connect_node
Input:
{
    "nodes" : ["http://127.0.0.1:5004"]
}
Output:
{
  "message": "All the nodes are now connected. The Blockchain
   now contains the following code",
  "total_nodes": [
    "127.0.0.1:5004",
    "127.0.0.1:5003",
    "127.0.0.1:5002"
  ]
}
```

```
Request Method Type: POST
http://127.0.0.1:5002/create_hash
Input:
{
    "file_path" : "MRI.pdf.aes"
}
Output:
{
    "IPFS Hash": {
        "Hash": "QmbWxSqxhZms74885xceSaiHVvdjqiVWrATtYUrrUnVE1K",
        "Name": "MRI.pdf.aes",
        "Size": "63884"
    }
}
```

```
Request Method Type: POST
http://127.0.0.1:5002/add_transaction
Input:
{
    "user_id" : "xxxxx",
    "ipfs_hash" : "QmbWxSqxhZms74885xceSaiHVvdjqiVWrATtYUrrUnVE1K",
    "EHR_name" : "MRI"

}
Output:
{
    "message": "This transaction will be added to blockchain 2"
}
```

```
Request Method Type: GET
http://127.0.0.1:5004/get_chain
{ "chain": [
{ "index": 1,
"nonce": 1,
"previous_hash": "0",
"timestamp": "2020-09-02 23:31:41.258101",
"transactions": [] },
{ "index": 2,
"nonce": 533,
"previous_hash": "03bee3d22f9c590f051f15f6efe33f282cfeac3fd90957779f04ddf29c2c879d",
"timestamp": "2020-09-02 23:33:07.304063",
"transactions": [
{ "EHR_name": "MRI",
"ipfs_hash": "QmbWxSqxhZms74885xceSaiHVvdjqiVWrATtYUrrUnVE1K",
"user_id": "xxxxx" } ] } ], "length": 2 }
```

```
Request Method Type: POST
http://127.0.0.1:5004/get_file_hash
Input:
{
    "name" : "MRI.pdf.aes"
}
Output:
{
    "File info": {
        "Hash": "QmbWxSqxhZms74885xceSaiHVvdjqiVWrATtYUrrUnVE1K",
        "Name": "MRI.pdf.aes",
        "Size": "63884"
    }
}
```

```
Request Method Type: POST
http://127.0.0.1:5004/get_file
Input:
{
    "ipfs_hash" : "QmbWxSqxhZms74885xceSaiHVvdjqiVWrATtYUrrUnVE1K"
}
Output:
{
    "The File has Retrieved at Your Local Directory"
}
```

Fig. 4 Necessary coding blocks for secure EMR preservation and sharing over the blockchain network

IPFS, by default, uses the local port http:/127.0.0.1:5001/, hence we kept port 5001 for the IPFS network. We considered port 5002 for the patient and port 5003 for the healthcare authority. However, we can add as many ports for individual identity, either a patient or a legitimate user of the EMR. The function used here is *connect_node* with **POST** method request.

Getting Hash from IPFS: After receiving the EMR from healthcare A, the patient first encrypt the it and then uploads in on the IPFS network to receive a SHA256 hash key from it Fig. 4b. At this stage, the function called *create_hash* has been used, and the request was **POST** method request. The input for the function was provided in JSON format, so the user only needs to provide the directory of the EMR. After running the function, the patient will get the unique hash key with other information like file name and file size in bytes.

Adding the EMR on the Blockchain: During the uploading stage, the patient needs to provide the file name of the EMR so that using that name he can query and retrieve the EMR from the blockchain when necessary Fig. 4c. In addition to EMR name, the patient also provides his name and the unique hash key received from IPFS. Here we have uploaded a sample MRI report received from healthcare A. The function used for the operation is *add_hash* with a *POST* request.

Getting the EMR Lists: Lists of all previous EMR can be found by using the function *get_chain* Fig. 4d with a **GET** request. By initiating the function the patient can easily get the lists of all previous EMR stored on the blockchain. We have shown the list with Genesys block and an MRI test report EMR. At the end of the EMR history, the length of blocks of the blockchain has been shown.

Getting the Specific EMR from Lists: When the patient needs to retrieve a specific EMR from the lists of EMRs, he just needs to run the **POST** request with function *get_file_hash* Fig. 4e. User needs to provide the name of the EMR in JSON format. The function then returns the file info by querying the whole lists of blocks.

Getting the Targeted EMR: Finally, to get the required EMR, user need to run the *get_file* function which a *POST* request Fig. 4f. When the user provides the unique IPFS hash key for the specific EMR, the function gets the file from the IPFS network and saved it into the local directory of the user. Finally, the user can share his private key to decrypt the document to the authorized personnel.

4.3 Performance Analysis

We have analyzed the performance of the proposed model in this section. The average latency of the system, during the execution of different functions, has shown in Fig. 5. The latency for the *mine_block* and *get_file* are higher in comparison with other functions.

Security Analysis: In the proposed system, if a user wish to access the medical reports of a patient, he need to get the private key from the owner of the data, as the data is encrypted using AES-256 algorithm. Besides, the full control of the data in under

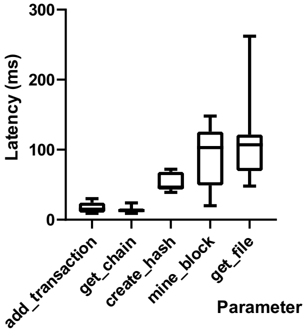


Fig. 5 Average latency of the proposed framework considering a MRI report of 62.3 KB

the control of the user, as the private blockchain was used in the proposed system. This was a unique achievement in comparison to related works mentioned in Sect. 2 [3, 27].

Scalability: The scalability in a simple term defined as the ability of the system to perform better irrespective of system data size increases or decreases. The scalability level of the proposed system is pretty good as the functions used here are lightweight and very low in size.

Integrity: The integrity of a system means the trustworthiness of the data store or transmitted in the system is unchangeable and reliable. Our system is highly reliable for patients as well for the stakeholder. The system ensures that data are immutable, using the decentralized and temper-proof nature of blockchain.

5 Conclusion and Future Work

In this article, we addressed how blockchain and IPFS technologies can be helpful to the healthcare industry, and how these can support handling of EMR. Despite advancements in the healthcare sector and technical advancement in EMR systems, they still faced some problems that this latest technology, i.e., blockchain, can tackle. Our suggested system incorporates the secure preservation of patient data with remote access guidelines for such data. It offers such a structure to promote users' usage and comprehension. The proposed system also concentrated on reducing document sizes for upload over the blockchain network. Moreover, unique IPFS hash and patient power over EMR guarantee data immutability.

We intend to introduce the platform for patients in certain local hospitals and online API. We would also attempt to simplify reimbursement for diagnostic testing and storage of EMRs. We should also define more user-based access guidelines and such strategies to follow healthcare legal expectations and values. We would also like to introduce a suitable healthcare data encryption system before sharing over a distributed network.

References

1. Al Omar, A., Rahman, M.S., Basu, A., Kiyomoto, S.: Medibchain: A blockchain based privacy preserving platform for healthcare data. In: Proceedings of SpaCCS, pp. 534–543. Springer (2017)
2. Arifeen, M.M., Al Mamun, A., Kaiser, M.S., Mahmud, M.: Blockchain-enable contact tracing for preserving user privacy during Covid-19 outbreak. Prrpints (2020)
3. Arifeen, M.M., et al.: Hidden Markov model based trust management model for underwater wireless sensor networks. In: Proceedings of ICCA, pp. 1–5 (2020)
4. Asif-Ur-Rahman, M., et al.: Toward a heterogeneous mist, fog, and cloud-based framework for the internet of healthcare things. IEEE Internet Things J. **6**(3), 4049–4062 (2018)

5. Azaria, A., Ekblaw, A., Vieira, T., Lippman, A.: Medrec: using blockchain for medical data access and permission management. In: OBD, pp. 25–30. IEEE (2016)
6. Benet, J.: Ipfs-content addressed, versioned, p2p file system. arXiv preprint arXiv:1407.3561 (2014)
7. Benet, J.: What is IPFS? (2020). https://docs.ipfs.io/concepts/what-is-ipfs/#decentralization. Last accessed 11 Aug. 2020
8. Biswas, S., Akhter, T., Kaiser, M., Mamun, S., et al.: Cloud based healthcare application architecture and electronic medical record mining: An integrated approach to improve healthcare system. In: 2014 17th International Conference on Computer and Information Technology (ICCIT), pp. 286–291. IEEE (2014)
9. Boonstra, A., Versluis, A., Vos, J.F.: Implementing electronic health records in hospitals: a systematic literature review. BMC Health Serv. Res. **14**(1), 370 (2014)
10. Chen, Y., Li, H., Li, K., Zhang, J.: An improved p2p file system scheme based on IPFS and blockchain. In: Big Data, pp. 2652–2657. IEEE (2017)
11. Daemen, J., Rijmen, V.: The Design of Rijndael: AES—The Advanced Encryption Standard. Springer (2002)
12. Dey, T., Jaiswal, S., Sunderkrishnan, S., Katre, N.: Healthsense: A medical use case of internet of things and blockchain. In: Proceedings of ICISS, pp. 486–491. IEEE (2017)
13. Dubovitskaya, A., Xu, Z., Ryu, S., Schumacher, M., Wang, F.: Secure and trustable electronic medical records sharing using blockchain. In: AMIA Annual Symposium Proceedings. vol. 2017, p. 650. American Medical Informatics Association (2017)
14. Esposito, C., De Santis, A., Tortora, G., Chang, H., Choo, K.K.R.: Blockchain: a panacea for healthcare cloud-based data security and privacy? IEEE Cloud Comput. **5**(1), 31–37 (2018)
15. Fazil, U.: A distributed file store (2020). https://medium.com/block360-labs/ipfs-a-distributed-file-store-533cda4c6047. Last accessed 22 Aug. 2020
16. Fazil, U.: Understanding AES 256 encryption (2020). https://www.solarwindsmsp.com/blog/aes-256-encryption-algorithm. Last accessed 02 Oct. 2020
17. Gordon, W.J., Catalini, C.: Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. Comput. Struct. Biotechnol. J. **16**, 224–230 (2018)
18. Gunter, T.D., Terry, N.P.: The emergence of national electronic health record architectures in the united states and Australia: models, costs, and questions. J. Med. Internet Res. **7**(1), e3 (2005)
19. Jenkins, J., et al.: Bio-mining for biomarkers with a multi-resolution block chain. In: Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XIII. vol. 9496, p. 94960N. ISOP (2015)
20. Juliana, D.G.: The history of data breaches (2020). https://digitalguardian.com/blog/history-data-breaches. Last Accessed 01 Sept. 2020
21. Kaplan, R.S., Porter, M.E.: How to solve the cost crisis in health care. Harv. Bus. Rev. **89**(9), 46–52 (2011)
22. Khan, M.I., Faisal, F., Azam, S., Karim, A., Shanmugam, B., De Boer, F.: Using blockchain technology for file synchronization. In: IOP Conference Series: Materials Science and Engineering, vol. 561, p. 012117. IOP Publishing (2019)
23. Kumar, R., Marchang, N., Tripathi, R.: Distributed off-chain storage of patient diagnostic reports in healthcare system using IPFS and blockchain. In: COMSNETS, pp. 1–5. IEEE (2020)
24. Mah'd Alloubani, A., Almatari, M., Almukhtar, M.M.: Effects of leadership styles on quality of services in healthcare. Eur. Sci. J. **10**(18) (2014)
25. Mahmud, M., Kaiser, M.S., Hussain, A.: Deep learning in mining biological data. arXiv (2020)
26. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. Technical report, Manubot (2019)
27. Rahman, A., Roy, S., Kaiser, M.S., Islam, M.S.: A lightweight multi-tier s-MQTT framework to secure communication between low-end IOT nodes. In: Proc. NSysS, pp. 1–6. IEEE (2018)
28. Rghioui, A.: Managing patient medical record using blockchain in developing countries: challenges and security issues. In: Morgeo, pp. 1–6. IEEE (2020)

29. Sears, K., Stockley, D.: Influencing the quality, risk and safety movement in healthcare: In: Conversation with International Leaders. CRC Press (2017)
30. Semantha, F.H., Azam, S., Yeo, K.C., Shanmugam, B.: A systematic literature review on privacy by design in the healthcare sector. Electronics **9**(3), 452 (2020)
31. Shahnaz, A., Qamar, U., Khalid, A.: Using blockchain for electronic health records. IEEE Access **7**, 147782–147795 (2019)
32. Tian, H., He, J., Ding, Y.: Medical data management on blockchain with privacy. J. Med. Syst. **43**(2), 26 (2019)
33. Wang, S., Zhang, Y., Zhang, Y.: A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems. IEEE Access **6**, 38437–38450 (2018)
34. Weisbaum, H.: The total cost of a data breach-including lost business-keeps growing (2020). https://rb.gy/r1njhr. Last accessed 01 Sept. 2020
35. Yue, X., Wang, H., Jin, D., Li, M., Jiang, W.: Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J. Med. Syst. **40**(10), 218 (2016)