IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 October 2023, accepted 12 November 2023, date of publication 14 November 2023,
date of current version 27 November 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3333229*

SURVEY

# Blockchain and Machine Learning in EHR Security: A Systematic Review

**UMER ZUKAIB[1], XIAOHUI CUI[1], MIR HASSAN[2], (Member, IEEE), SHEETAL HARRIS[1], HASSAN JALIL HADI[1], AND CHENGLIANG ZHENG[1]**

[1]Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China
[2]Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy

Corresponding author: Xiaohui Cui (xcui@whu.edu.cn)

This research was supported by Research and Application of Object detection based on Artificial Intelligence of Science and Technology Department of Yunnan Province (No.2022KZ00125).

**ABSTRACT** Background: The rapid development of modern technologies renders a convenient and efficient solution to implement Electronic Health Records (EHRs) systems. The rapid growth of healthcare data has a distinctive attribute of digital transformations. The big datasets of healthcare, their complexity and their dynamic nature have posed severe challenges associated with the analysis, pre-processing, privacy, security, storage, usability and data exchange. Material and Methods: We have performed the Systematic Literature Review (SLR) and followed the Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) methodology. SLR refers to the methodology that discovers, analyses and accesses recent research literature related to the subject field. The research papers were searched from academic repositories like IEEE, WOS, Scopus and PubMed for the previous five years on March 2023. Results: The designed search string provides 199 research articles in total. We filter the research articles based on inclusion-exclusion strategies and quality assessment metrics. Six main criteria for research inclusion-exclusion for SLR are formulated. These works of literature insight into 1) the issues associated with interoperability and security of EHRs by using the Blockchain (BC) technology, 2) different frameworks and tools to improve privacy and security in the healthcare domain, 3) the open issues of using BC technology in the electronic healthcare domain, 4) the standardized ways to store EHRs, 5) various ways to handle the big data using the BC systems and 6) the usage of Federated Learning (FL) to preserve the privacy of EHRs in the healthcare domain. We acquired 46 research articles based on the criteria (inclusion-exclusion) that investigate the above-mentioned issues. Conclusion: The SLR will serve as the state-of-the-art (SOTA) for future researchers in the field of BC in healthcare. Additionally, the paper provides insights to the new researchers to revolutionize the healthcare domain by adopting the latest digitalized technologies. The proposed study identified various reflections. It analyzed the architectural mechanism that supports the security and interoperability of EHRs. Secondly, the study described different tools and frameworks to improve the privacy and security of EHRs using the BC. Thirdly, the open issues of storing and preserving the EHRs using BC in the healthcare system were determined. Fourth, it analyzed and provided a detailed view of using standardized ways for storing and handling big data by using the BC system. Lastly, the usage of FL to preserve the privacy of EHRs was analyzed.

**INDEX TERMS** Blockchain technology, smart healthcare, federated learning, securing patient records, deep learning, electronic health records.

## I. INTRODUCTION

In the digitalized world, global healthcare coverage is 50%. Everyone needs to access quality healthcare services (disease

The associate editor coordinating the review of this manuscript and approving it for publication was Tyson Brooks.

© 2023 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

diagnosis, prevention and treatment) in a convenient, safe, transparent and efficient way [1]. Keeping the purpose in view, the technologies that provide coverage, easy access, and boost the quality of health services are continuously working to achieve perfection. In the absence of quality health services, medical centres might become inefficient and lose their credibility [2]. The larger healthcare organizations have many inter-connected stakeholders that have different competing interests. The healthcare ecosystem involves a comprehensive, trusted and reliable Personal Identifiable Information (PII) exchange among various stakeholders [3]. This PII has been distributed and fragmented into multiple non-integrated data-storage systems that hinder information access. Thus, the decision-making process gets affected. The reason is that each hospital or medical centre manages healthcare data in a centralized manner, causing the health personnel to have a brief patient history, which leads to errors in the diagnosis and treatment process. Centralized information leads to different information risks. As highlighted by [4], the healthcare domain is the most exposed to cyber-attacks, like ransomware or Denial-of-Service attack (DoS) [5], which hijacks the PII, and in most cases, it is impossible to recover back the information that exposed to cyber-attack [6]. For these reasons, a robust mechanism is required to achieve interoperability among information systems which support the care process [7].

Recently, healthcare has been one of the crucial concerns. Every second massive amounts of data are generated, stored and used frequently. EHRs are a subset of the healthcare system. It allows the medical history to be accessed by the doctors, patients and other medical staff to avoid repetitive imaging, radiological and other expensive tests [8]. The main challenge of EHRs remains preserving the patient's privacy. Accessing the patients' records with utmost privacy is a significant challenge. The second important issue for EHRs is that the patients do not own their data. Hence, medical centres own the patient's medical data. The access to the patients' EHRs without their consent and use this data for the purpose other than research and treatment is one of the challenging aspect of patient privacy [9]. EHRs have different challenges in terms of security, such as the frequent use of IoT and wearable sensor devices to diagnose diseases and store the data in the medical file of the patient. This approach has a high-security threat and has a maximum risk of attack. The physician's prescription for curing disease can be compromised, which endangers the patient's life [10]. Another security issue is fraud detection. There have been many cases in which doctors gave prescriptions for drugs that were not necessary for the patient. The mere purpose of prescribing the drugs is the availability at the medical store of the same hospital where the doctor works. As a result, the patient's health is compromised, along with unnecessary medical expenditures [11]. Another security risk is counterfeit medicines. To address this issue, the drug supply chain is arranged such that critical information is easily accessible. It also contains information regarding the pharmaceutical plant that produces the drugs along with the storage, transportation and distribution details [12].

Additionally, security is an essential element for distributing patients' EHRs, and it also has serious risks associated with the healthcare system that cause damage to the finance, insurance, reputation and some other essential factors of the healthcare ecosystem [13]. To resolve all these issues, BC technology is used. BC is the combination of decentralization, consensus mechanisms, cryptographic hashing, and the widespread distribution of copies of the ledger ensures that once data is added to the BC, it becomes practically impossible to change or tamper with. This immutability is a core feature of BC technology and is essential for its trustworthiness in various applications, including finance, supply chain, and healthcare [14].

The healthcare domain is constantly developing with the introduction of new tools and approaches. To improve the integrity of medical data, confidentiality, traceability, security and interoperability, BC technology is used. There is a relationship between the BC and the management of EHRs [15], [16]. Besides the security risk of distributing EHRs, another issue is associated with data delivery and management. For the safe delivery and management of EHRs cloud computing plays a crucial role by enables the secure distribution of EHRs and provide a flexible and scalable infrastructure for the healthcare organizations. EHRs are encrypted when stored in the cloud, ensuring that even if unauthorized access occurs, the data remains unreadable. Cloud platforms enable healthcare organizations to implement granular access controls. Cloud providers offer Identity and Access Management (IAM) services that facilitate user authentication and authorization, ensuring that only authorized personnel can access the patient records. Multi-Factor Authentication (MFA) add the extra security layer, essential for all the users to provide multiple forms of verification before gaining access to EHRs [17].

Deep learning (DL) also play an important role in enhancing the security of EHRs within the healthcare organizations. While DL models are often associated with various applications in healthcare, they can also contribute to EHRs security. DL models observe the anomaly behavior or patterns of access in EHRs, especially the auto-encoders and recurrent neural networks (RNNs) can improve overall security by triggering alarms when anomalous behaviors, unauthorized access or data breaches may occur [18]. DL techniques can facilitate the methods of biometric authentication, like fingerprint scanning, facial recognition and authorized EHR access. To analyze and process the text data within EHRs, DL based NLP models provide the most appropriate solutions. NLP models identify and categorize the sensitive information and help to ensure proper handling, redaction of confidential patient's data to protect the patient privacy [19]. Besides DL model perform the task of secure medical image-processing, perform data-encryption

and optimization, predictive analysis, network security like intrusion-detection systems and continuous authentication. It is essential to integrate DL models effectively within a comprehensive security framework to regulate the update and adapt security measures to evolving threats, ensuring that patient data remains protected and confidential [20].

Federated learning (FL) is an efficient approach for securing EHRs within healthcare organizations. It offers several advantages in terms of data privacy, security, and collaborative model training while preserving patient confidentiality. In FL, EHRs data remains on the local servers of each healthcare organization. FL allows the healthcare organizations to collaboratively train the machine learning models without sharing the raw EHRs data. Only the model updates are shared, aggregating model updates from multiple organizations which are aggregated to improve the global model while keeping the data decentralized and secure. This aggregation is done in a way that prevents the individual records from being exposed and also maintaining the anonymity of the patients. FL is an efficient and secure approach for EHRs in healthcare organizations, also strikes a balance between data privacy, security, and collaborative model training, ensuring that sensitive patient information's are protected while still allowing for valuable insights to be derived from the data. FL approach aligns with healthcare data privacy regulations and helps maintain patient trust in the healthcare system [21].

With the recent evolution of BC technology and healthcare systems, it is difficult to figure out state-of-the-art (SOTA). Keeping this in view, we propose a Systematic Literature Review (SLR). We adopt the PRISMA technique and find out the architectural mechanisms that are used to enhance interoperability and security of EHRs by using BC, different frameworks and tools to improve the privacy and security in the healthcare domain, the open issues of using BC technology in the electronic healthcare domain, the standardized way to store EHRs, handle big data using the BC systems and finally the usage of Federated Learning (FL) to preserves the privacy of EHRs in the healthcare domain.

### A. CONTRIBUTION

This section outlines the research questions (RQs) that guide the systematic literature review (SLR) and provides an overview of the methodology used to address these questions, highlighting the importance of the review in the context of electronic health records (EHRs) and Blockchain technology.

*Research Questions (RQs):*

1) What architectural mechanisms have been used to support the security and interoperability of electronic health records using Blockchain technology?
2) What are the different frameworks and tools used to improve privacy and security in healthcare, specifically for securing electronic health records using Blockchain technology?
3) What are the open issues related to the usage of Blockchain technology in the electronic healthcare domain and their future perspectives?
4) What is the standardized way to store Electronic Health Records (EHRs) and handle big data in Blockchain systems?
5) How does Federated Learning (FL) preserve the privacy of EHRs in the healthcare domain?

These research questions address the need for conducting this Systematic Literature Review (SLR) by:

- Designing a search string based on keywords to efficiently search different databases for articles relevant to the research topic.
- Following the PRISMA methodology for conducting the SLR.
- Performing the literature search from 2018 to 2023 using the unique search string in various databases.
- Defining initial quality metrics for the initial review of articles, based on four unique questions.
- Conducting a quality assessment of articles using six questions, assigning a quality score to each article, and including those with a quality score of 50 % or higher.
- Evaluating research articles based on inclusion and exclusion policies, consisting of five questions for each.
- Providing answers to the designed research questions (RQ1 - RQ5) based on the scrutinized literature.

This manuscript is structured as follows: In Section II we provide an overview of the existing literature to contextualize our research. Section III, outlines the review methodology employed, detailing the systematic approach used to answer the research questions. Section IV, presents the findings and answers to the research questions. Section V, discusses the challenges encountered during the study and the corresponding technical solutions. Section VI, addresses limitations and suggests potential directions for future research. Finally, in Section VII, we summarize our findings and contributions.

## II. RELATED WORK

### A. BACKGROUND

BC technology uses a distributed, decentralized ledger to log transactions via a computer network. BC is utilized in the medical field for a number of reasons. BC offers a secure and immutable way to store EHRs and patient data. Every transaction (like adding or modifying a patient record) is tracked as a "block" in the chain, which is then connected and safeguarded by cryptographic hashes. This guarantees that information cannot be changed or tampered with without leaving a trail. BC technology is able to control EHRs access. With smart contracts, patients can manage who has access to their data, allowing and removing rights as necessary. This improves consent management and patient privacy. BC can help with interoperability by offering a standard framework for exchanging EHRs among various healthcare organizations and systems. While preserving data security, smart contracts can automate the process of exchanging data between parties. BC ensures the authenticity and quality

**TABLE 1. Abbreviations and descriptions.**

| Abbreviation | Description |
|---|---|
| SLR | Systematic Literature Review |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| IoT | Internet of Things |
| BC | Blockchain Technology |
| FL | Federated Learning |
| ML | Machine Learning |
| DL | Deep Learning |
| AI | Artificial Intelligence |
| EHR | Electronic Health Records |
| SOTA | State of the Art |
| PII | Personal Identifiable Information |
| IAM | Identity and Access Management |
| NLP | Natural Language Processing |
| RNNs | Recurrent Neural Networks |
| RQs | Research Questions |
| BVOABSC | Attribute Based Signcryption Scheme |
| DST | Discrete Shearlet Transform |
| CHG | Cryptographic Hash Generator |
| PCA | Principal Component Analysis |
| QA | Quality Assessments |
| IQ | Initial Quality |
| FHIR | Fast Health Interoperability Resources |
| DoS | Denial-of-Service |
| IoMT | Internet of Medical Things |
| SMPC | Secure Multi-Party Computation |
| HE | Homomorphic Encryption |
| DSL | Domain Specific Language |
| API | Application Programming Interface |
| CIM | Clinical Information Model |
| MST | Multiple subpial Transaction |
| DP | Differential Privacy |
| SGD | Stochastic Gradient Descent |
| VMR | Virtual Medical Records |
| HIPAA | Health Insurance Portability and Accountability Act |
| CISA | Cyersecurity Infrastructure Agency |

of products by tracking the provenance of pharmaceuticals and medical devices in the supply chain. BC can streamline clinical trial management by securely recording and sharing trial data, ensuring transparency and data integrity [22], [23].

FL is a machine learning technique that enables several parties to work together without disclosing their raw data to train a common machine learning model. FL has the following advantages for the healthcare industry. Healthcare organizations, such as hospitals and research institutions, can keep sensitive patient data on their own premises. Instead of sending data to a centralized location, FL sends model updates or gradients to a central server. This ensures that patient data remains within the organization's control, enhancing privacy and security. Without breaking data privacy laws, researchers and medical practitioners can work together to jointly train ML models utilizing data from various sources. FL preserves patient data security while facilitating the sharing of insights. In FL, the central server aggregates the model updates from different participants. It then uses these updates to improve the global model. This process ensures that the shared model becomes more accurate without exposing sensitive data. FL facilitates real-time and ongoing model refinement. The model can be improved and updated at the participant sites when new data becomes available. FL can scale to include numerous healthcare institutions, accommodating the sharing of data and model updates in large, distributed networks [24].

### B. BLOCKCHAIN (BC) IN HEALTHCARE

In the scenario of smart health, the most crucial problem is the security of EHRs. The challenging task is securing the smart health system and reducing and encrypting the data by compression and encryption models. BC system consists of various stakeholders like hospitals, pathologists, pharmacists, physicians, insurance, and researchers. Benil et al. [25] described that BC provides a convenient solution to address EHRs storage and efficient retrieval limitations. Usually, in BC, the patients' EHRs are encrypted by patients' private keys and stored in a ledger. BC is not completely impenetrable, and it is more secure compared to ordinary systems. Razzaq et al. [26] suggested that the EHRs data leakage enhances the patient's privacy risk. Generally, the EHRs data is unchanged after being uploaded to the system. The data stored in the BC ledger can be shared with all the medical parties and stakeholders. The study proposed BC based cryptographic technique. The EHRs indexing was stored in the BC, and because of this, patients can access and view EHRs data. In their proposed system, only the search indices were stored in BC, which facilitates the distribution of EHRs. In comparison, the original encrypted EHRs data was stored on another server. Shah et al. [27] solved the problem of data collaboration and deployed healthcare systems in a cloud-based heterogeneous environment. The researcher also used the ChainSDI framework, which is computationally efficient. The implemented prototype showed the framework had securely managed all the data. Karimi et al. [28] used containers in the BC framework to secure healthcare data. For the improvement of data transferring, the container was connected with multiple ports. Besides this, a medichain framework based on the BC platform was also proposed in the same research study. Separately in each block, the EHRs of patients were maintained, and the security features of BC secured the health data. Python language and object-oriented concepts were used to implement the framework prototype.

Jayabalan et al. [29] emphasized securing and protecting the patient's EHRs from unauthorized intrusion. First, they figured out how to control and manage the data access, and then BC and cryptographic algorithms were implemented for secure data storage and transmission. Before the data storage and transfer, the cryptographic algorithm encrypts the data. The simulation results of the proposed scheme demonstrated efficient data transferring and secure data storage compared to existing schemes. Wang et al. [9] emphasized the security and privacy issues of the healthcare domain. To provide more security to healthcare data, they focused on important features of BC like zero-knowledge proof, smart contracts approvals, attribute-based encryption,

and anonymous signatures. Secure data sharing was achieved using several security techniques. Parekh et al. [30] discussed Patient Remote Monitoring (PRM), and BC-based architecture was proposed for the effective storage and transferring of healthcare data. Chenthara et al. [31] used smart contracts for proper management and analysis of medical data. Besides this, smart contracts also analyze the data generated by the sensors attached to the body of patients. If a patient's data were in danger of cyber-attack, a warning message is triggered to the medical center for intensive care.

Qiu et al. [32] proposed BC based data-management framework and used asymmetric encryption to secure the system's data. It enables patients to access medical records from various remote locations. Lee et al. [29] proposed BC and a cloud-based system that efficiently stores and shares EHRs. The authors identified challenging issues in the current healthcare domain and implemented a prototype that effectively solved the existing issues. For testing the prototype, they used Amazon and Ethereum-based BC systems, and IPFS was used for data storing and sharing. The results demonstrated the proposed system efficiently stored and shared medical information. The proposed system detects any type of unauthorized access to the data and prevents data theft. The system showed BC is an effective solution for managing EHRs.

BC has extensive and attractive healthcare applications commonly used to secure the transmission of EHRs, and medical data or maintain the medical supply chain [33]. Al-Sumaidaee et al. [34] proposed a public BC-based system that guarantees the validity and confidentiality to handle a single failure point by managing EMRs by fixing data encryption and scalability issues. Abouali et al. [35] proposed a private BC system that provides access control and a privacy-based solution. The data was collected using an Android application from different sources such as medical teams, patient records, and insurance companies. The proposed solution also addressed the scalability and efficiency problems associated with data processing.

Rehman et al. [36] proposed an RTS-DELM system based on BC that secures the patient records. FL provides a facility to transfer EMD securely between hospitals and IDS that can detect malicious activities on healthcare networks, and the model also performs the task of disease prediction. The proposed model scored 96.18 % accuracy for intrusion detection over the network and predicted diseases with 93 % accuracy. Yang et al. [37] proposed a Blockchain and Attribute-based Signcryption-Scheme (BVOABSC) that provides flexible control and access to EHRs on the cloud servers and reduces the storage cost and bandwidth utilization. The scheme reduces the computational burden of the users by allowing cloud servers to perform all the decryption operations. The ciphertext produced by the cloud server is validated by users. It demonstrated that the proposed scheme ensures verifiability. As compared to existing techniques, ABVOABSC has high computing power and low communication overhead. Hoang et al. [38] proposed a HoloCare system based on BC that provides a secure channel for sharing and remotely accessing Personal Health Records (PHR). The system was implemented in hyper-ledger fabrics with a CFT consensus algorithm. The performance was evaluated by the hyper ledger caliper tool. The tool creates 10,000 transactions of query and invokes types to the BC network. The invoke transaction model scored a min latency of 0.06, and the query transaction model scored a min latency of 0.01.

Hegde et al. [39] adopted three off-chain methods (IPFS, CosmoDB and Storj) for storing the EHR and experimented with faster storage and retrieval time on different off-chain methods. The study was performed on the Chronic kidney disease dataset, and the XGBoost model was used for disease prediction. The results showed that IPFS performed well. The average time for IPFS data storage is 0.11s, and 0.095s for data access. Marichamy et al. [40] proposed BC and Hadoop-based systems for securing medical records and used a Cryptographic Hash-Generator (CHG) to secure the user key. Before storing medical data in BC, the data is encrypted using Discrete-Shearlet Transform (DST) for verification purposes. CHG generates the request (sent by the user), and the operator creates a block using the remote key and triggers a request to the requester to sign and fulfill the order of request using the private key. For verification of optimal-request from the queue, the Improved Grey-Wolf Optimization algorithm (IGWO) is used. The model is evaluated on several parameters. At 10Mb of data, the model utilized 22.91 % of memory for data encryption. Download-time for 10Mb of data was 4.61s, and upload-time was 7.9s. The model showed 99.0% accuracy on 10 MB of data, and the lower-attack level was 12.5% on 10MB of data. The proposed system is compared with other models, and the proposed system outperformed the other models. Some well-known research studies based on blockchain and deep learning, along with blockchain and machine learning in healthcare domain are presented in Table 2 and Table 3.

#### 1) TECHNICAL ASPECTS OF BLOCKCHAIN IN HEALTHCARE DOMAIN

Blockchain technology is useful for the healthcare industry since it provides a number of technical information and functionalities. Some important technical features of blockchain in healthcare are as follows:

- **Decentralization:** Since Blockchain is a decentralized ledger, it records and validates transactions without the help of a central authority. By doing away with the necessity for middlemen, decentralization in healthcare can increase productivity and lower the possibility of data manipulation.
- **Distributed Ledger:** A network of nodes, or computers, makes up the distributed ledger of the blockchain. Since each member of the network has a copy of the ledger, there is less chance of a single point of failure, and data is widely distributed.

- **Immutability:** Data cannot be changed or removed once it is stored on the blockchain. The integrity of medical records and other vital healthcare data is guaranteed by this immutability. New blocks are added to the ledger whenever there are modifications or additions.
- **Security:** To protect data, blockchain uses robust cryptography algorithms. This contains consensus techniques that stop unwanted changes, encryption of data while it's in transit and at rest, and public and private keys for user authentication.
- **Consensus Mechanisms:** To validate and append new transactions to the ledger, blockchains employ consensus algorithms. Consensus algorithms such as Proof of Work (PoW) or Proof of Stake (PoS) guarantee that data is uploaded only after agreement from several parties, which is important in the healthcare industry where data veracity is crucial.
- **Smart contracts:** Smart contracts are self-executing agreements that have the provisions of the contract explicitly encoded into the code. Smart contracts can reduce the need for middlemen in the healthcare industry by automating procedures like supply chain tracking, insurance claim processing, and patient consent management.
- **Blockchains with permissions:** Healthcare companies frequently employ blockchains with permissions, which restrict network access to authorized users exclusively. This guarantees the privacy of important patient information.
- **Data Interoperability:** Blockchain technology can act as a standard framework to facilitate interoperability across different EHRs and healthcare systems. It permits various organizations and entities to have control over their own data while securely accessing and exchanging data.
- **Data Provenance:** Blockchain preserves data history, offering an auditable and transparent trail of data modifications and accesses. This function is essential for monitoring how healthcare data is used and shared as well as for guaranteeing legal compliance.
- **Privacy and Confidentiality:** Transactions can be verified without disclosing the real data by using strategies like zero-knowledge proofs. This preserves patient confidentiality while guaranteeing the authenticity and accuracy of the data.
- **Scalability:** In some blockchain systems, especially public blockchains, scalability is a hurdle. However a number of strategies, like sharding and off-chain solutions, are being investigated to solve this problem and make blockchain capable of managing a higher number of healthcare transactions.
- **Data standardization:** Healthcare blockchains frequently follow defined data formats, like Fast Healthcare Interoperability Resources (FHIR), to facilitate smooth interaction with current healthcare systems and guarantee efficient data sharing.
- **Data Ownership and Consent Management:** Patients now have more influence over data ownership and consent management related to their healthcare thanks to blockchain technology. They have the ability to transparently see who has seen their data and to grant and withdraw access to their records.

These technical aspects of blockchain in healthcare improve the ecosystem's efficiency, transparency, data security, and interoperability. To guarantee that blockchain technology is used effectively and patient data is treated responsibly, it is crucial to address regulatory compliance, data governance, and ethical issues.

### C. DEEP LEARNING (DL) IN HEALTHCARE

DL is the branch of AI that provides an analytical model by automating and analyzing the data. ML and DL algorithms learn from the data patterns and make decisions with minimalistic human intervention. Advanced ML, such as the automated ML field, has automated the model's building process from the pre-processing phase to the evaluation phase [18], [50]. Usually, traditional software is based on program code (rules for the system's behaviour), but in ML, the rules are deducted from the training data using an ML algorithm. ML algorithms create the rules according to the data types [12]. ML/DL is adopted as a supplementary technology in other emerging technologies like IoT and BC. For example, ML/DL with BC has numerous applications in the healthcare domain such as RPM, EHRs management, medical-data analysis, bio-medical study, supply chain, education, etc. BC has revolutionized modern technologies as it is integrated into the vast system, especially in bio-medical and healthcare systems [51]. The integration of BC and ML in IoT plays a crucial role in Industry 4.0 and IoMT [52].

DL models detect the data patterns associated with healthy or diseased conditions. The DL models are trained on historical datasets of medical or healthcare records. Recently ML and DL have been adopted to assist healthcare systems in developing innovative-sustainable solutions for treating chronic diseases like tumour diagnosis and its treatment [53]. DL and ML models are efficient for complex hidden data pattern classification. Therefore, ML and DL are commonly used in medical applications. Specifically, the applications depend upon genomics, proteomic analysis, and disease prediction.

Nancy et al. [19] proposed a DL-based heart disease-risk prediction system. The system comprised four layers, IoT layer, data acquisition, cloud layer, and final prediction layer. The IoT layer collects data and sends it to the data acquisition layer that stacks all the data. For pre-processing, data is forwarded to the cloud layer. The cloud layer has two sub-components data filtering and information system (FIS). For data filtering, a low-computational unsupervised Kalman Filter is used that removes useless data values. Next, the data is transferred to FIS, and finally, a Bi-LSTM model performs the prediction. The evaluation was performed based on accuracy, precision, and recall. The proposed system scored

**TABLE 2.** Research studies based on DL and BC in healthcare.

| Author | Objective | Area | Blockchain | Deep Learning | Accuracy |
|---|---|---|---|---|---|
| G. Zhang et al., [41] | Proposed BDL-SMDT network for detecting disease from medical images | BC and DL in Healthcare | BC is used to store encrypted-images | RESNet-V2 is used for feature-extraction and SVM is used for classification | 95.2% accuracy 96.94% sensitivity and 98.36% specificity |
| T. Veeramakali et al., [42] | Proposed Deep-learning as-a-service (DaaS) to store EHRs and forecast illness | Securing EHRs, DL and BC in Healthcare | BC is used to secure EHRs | LSTM is applied to forecast illness | 72.4% precision and 71.1% F1-score |
| S. Singh et al., [43] | Proposed a DL based secure BC enabled intelligent IoT healthcare diagnostic network (ODLSB) | Healthcare | For the hash value-encryption is used neighborhood-indexing-sequencing algorithm (HVE-NIS) | Optimal DNN model is applied to identify patient's disorder | 93% accuracy, 92% sensitivity and 91% specificity |
| P. Kumar et al., [44] | Design an efficient data-sharing and secure model by combing DL and BC | Healthcare | The participating entities have registered into BC ledger and confirmation was made by using zero knowledge-proof, and smart contracts were deployed for validation | Bi-LSTM and Auto-encoders is used for prediction | Bi-LSTM scored 99.58% accuracy and Auto-encoders score 99.98% accuracy |
| B. Mallikarjuna et al., [45] | Combine DL with BC and homomorphic-encryption | Healthcare | BC based IDS is developed | Used DL for prediction of security threats | 90.6% accuracy |
| G. Nguyen et al., [46] | Proposed a network by combines FL and BC that add security to smart health systems | Healthcare | BC ensure privacy and security | FL is applied for global training | NA |
| F. Ni et al., [47] | Proposed a deep-belief network that identifies intrusions from the sensors collected data | Healthcare | BC is used to provide data security and then data is shifted on the cloud server | ResNet is used for the classification of disease | 98.94% accuracy, 98.92% precision, 98.81% recall |
| V. Hassija et al., [48] | BC and DL-based network is used to analyze COVID-19 | Healthcare | BC is used to secure data | DL based feature-extraction technique is used to extract data | NA |
| A. Ali et al., [49] | Proposed a novel framework by combining DL and BC that identifies medical frauds in the insurance companies | Identification of medical-frauds | Consortium BC is used for storing and accessing EHRs | DL model is used to identify the medical-frauds | 90% accuracy |

98.86% accuracy, 98.90% precision, and 98.81% recall. The comparison analysis with previous literature showed the highest performance.

Khan et al. [59] proposed IoT-based CNN network, which is connected to wearable sensors for measuring blood pressure and monitoring the ECG of the patients. The proposed system is compared with other DNN models and logistic regression. Results demonstrated that the proposed CNN outperformed by scoring 98.2% of accuracy. An ensemble DNN Logistboost with a feature-fusion-based smart system was proposed by Ali et al. [60] that gathered data from wearable sensor devices. Based on the patient's medical history, it predicts the risk of heart disease. The system scored 98.5% accuracy for diagnosing heart disease and suggested a dietary plan for critical health situations. Zhang et al. [61] combined DNN with LinearSVC algorithm with an embedded feature-selection technique and tested the proposed model on the heart-disease dataset for predicting heart disease. The proposed model achieved 98.56% accuracy, 99.35% recall, 97.84% precision and 98.3% F-measures.

Gupta, et al. [62] proposed RF and ET-based techniques with FAMD-based feature extraction that extracts 96 features from the Z-Alizedeh Sani dataset. SMOTE was used to handle the data imbalance, and a 3:1 hold-out validation scheme was adopted. After parameter optimization, the model was evaluated based on accuracy, specificity, and sensitivity. By using binary PSO optimization, the proposed model scored 95.88% AUC, 97.37% accuracy, 95.45% specificity, and 98.15% sensitivity. Ghasemieh, et al. [56] proposed Stacking-Ensemble-Learner (SEL) with behavior-based features, which creates a new class label. The model was applied to the computational physiology dataset proposed by MIT. The model performs prediction of re-admission of heart-failure patients. The model was evaluated based on accuracy, recall, precision, and F1-score. Experimental results showed the ensemble model scored 88.23% accuracy and exhibited the best performance compared to other baseline models.

Kavitha et al. [54] proposed ML techniques for predicting Alzheimer's disease based on OASIS data. Three feature-selection techniques were used chi-square, information gain, and correlation coefficient. Experimental results demonstrated RF scored 86.92% accuracy, SVM scored 81.67% accuracy, DT scored 80.46% accuracy, XGBoost 85.92% accuracy, and the Voting classifier scored 85.12% accuracy. Krishnamoorthi et al. [63] proposed a framework

**TABLE 3.** Research studies based on ML and BC in healthcare.

| Author | Proposed Work | Dataset | Comparative Models | Proposed Model | Accuracy |
|---|---|---|---|---|---|
| N. Nissa et al., [20] | Proposed model for predicting the heart failure | Cardiovascular Disease dataset | RF, SVM and DT | RF | 99.35% |
| C. Kavitha et al., [54] | Proposed model for predicting | Alzheimer's disease dataset | XGBoost, DT, SVM and RF | XGBoost | 85.92% |
| S. Geetha et al., [20] | Proposed neural-network-based framework for estimating cardiac infection | Cleveland dataset | AdaBoost, SVM, DT, NB, and KNN | DT | 99.7% |
| P. Ghosh et al., [55] | Proposed ML model that predicts coronary-artery disease | Gather data from Cleveland, Switzerland and Hungary | AdaBoost, DT, GB and KNN | RF | 99.05% |
| A. Ghasemieh et al., [56] | proposed Stacking-Ensemble-Learner (SEL) that predicts heart-risk | computational-physiology dataset | SEL, RF and DT | SEL | 88.23% |
| I. Sardar et al., [57] | Proposed GLMNet for predicting COVID-19 cases | COVID-19 dataset | ARIMA, RF GLMNet, and XGBoost | GLMNet | 97.35% |
| Z. Ibrahim et al ., [58] | Proposed ANN-based ensemble approach to predict COVID-19 positive cases | COVID-19 dataset | ANN, MLR, ANFIS, SVM | Ensemble ANN | 0.96% R2 0.0002% MSE |
| M A. Khan et al., [59] | proposed IoT based CNN network to measure BP and monitor ECG | Public Health Dataset | CNN, DNN and LR | CNN | 98.2% |

called the Intelligent Diabetes-Mellitus-Prediction Framework (IDMPF) based on different ML techniques. Using the Indian diabetes dataset, Grid and random search were adopted for the HPO of models, and evaluation was made based on accuracy, ROC, test score and precision. The proposed model scored 83% accuracy. De Bois. et al. [64] adopted RETAIN framework to forecast the glucose level in diabetic patients and perform experiments on two datasets, IDIAB and OhioT1DM dataset. The performance was evaluated based on RMSE and MAPE, and RETAIN model outperformed other SOTA models.

### 1) TECHNICAL ASPECTS OF MACHINE LEARNING / DEEP LEARNING IN HEALTHCARE DOMAIN

The healthcare industry benefits greatly from machine learning and deep learning as they facilitate the study of big datasets, predictive modeling, individualized treatment suggestions, and diagnostic help. The following are some technical specifics of deep learning and machine learning in healthcare:

1) *Preprocessing of Data:*
   - *Data Cleaning:* Healthcare datasets frequently need to be thoroughly cleaned before training models in order to get rid of outliers, missing values, and inconsistencies that can have an impact on the model's performance.
   - *Feature engineering:* To extract valuable information from unprocessed healthcare data, domain-specific feature engineering is essential. Features could, for instance, reflect an image's texture, shape, or intensity in medical imaging.
2) *Supervised Learning:*
   - *Classification:* Machine learning models are applied to tasks such as forecasting patient outcomes, detecting anomalies in patient data, and classifying diseases (e.g., cancer detection).
   - *Regression:* Regression models are used to forecast numerical data, such as the length of hospital stays or the estimate of a disease's course.
3) *Unsupervised Learning:*
   - *Clustring:*Techniques for unsupervised learning can be used to find hidden patterns in medical data. Cohort analysis and patient stratification can both benefit from clustering.
   - *Dimensionality Reduction:* Data can be made easier to work with and visualize by reducing its dimensionality through methods like Principal Component Analysis (PCA).
4) *NLP, or Natural Language Processing:*
   - *Tokenization and Text Cleaning:* NLP models carry out lemmatization or stemming, tokenize text data into words or phrases, and eliminate stop words.
   - *Named Entity Recognition (NER):* Names, drug names, and medical conditions are just a few examples of the things that NLP models may extract from clinical notes.
   - *Sentiment Analysis:* Sentiment analysis is a useful tool for assessing patient opinions and feelings as they are described in narratives about healthcare.

5) *Deep Learning:*
   - *Convolutional Neural Networks (CNNs):* CNNs are utilized for image analysis tasks like pathology slide analysis, disease diagnosis in MRIs and X-rays, and medical image segmentation.
   - *Recurrent Neural Networks (RNNs):* RNNs are used for sequence data processing, including electrocardiogram (ECG) and vital sign time series analysis.
6) *Long Short-Term Memory (LSTM) Networks:* LSTMs are a particular kind of RNN that perform particularly well in applications where the ability to recall previous inputs is essential, like real-time patient deterioration prediction.
7) *Ensemble Methods:* By merging the results of several machine learning models, methods like Random Forests and Gradient Boosting can enhance model performance and produce predictions that are more accurate.
8) *Hyperparameter tuning:* Fine-tuning model hyperparameters is a critical technical aspect of optimizing model performance. Techniques like grid search, random search, and Bayesian optimization are used to find the best hyperparameters for a given task.
9) *Data augmentation:* To increase the training dataset's artificial size and enhance the generalization and robustness of models, data augmentation techniques are used in medical imaging.
10) *Model Interpretability:* Black-box machine learning models are interpreted using methods like SHAP (SHapley Additive explanations) and LIME (Local Interpretable Model-agnostic Explanations) to help healthcare practitioners better grasp the models' conclusions.
11) *Privacy-Preserving Techniques:* In the field of healthcare, where maintaining data privacy is of utmost importance, methods such as Homomorphic Encryption and Federated Learning enable machine learning models to be trained on encrypted, dispersed data without disclosing private information.
12) *Validation and Evaluation Metrics:* To properly validate and evaluate machine learning models, performance on healthcare tasks is measured using metrics such as accuracy, precision, recall, F1-score, AUC-ROC, and AUC-PR.
13) *Model Deployment:* When implementing machine learning models in the healthcare industry, it's important to take into account how to integrate them into clinical workflows, ensure their scalability, and adhere to legal requirements like HIPAA.

In the healthcare industry, where precise forecasts and insights can have a substantial impact on patient care, diagnosis, and treatment decisions, these technical details are crucial for the creation, implementation, and optimization of machine and deep learning models.

### D. FEDERATED LEARNING (FL) IN HEALTHCARE

In 2016 Google introduced FL to solve the issue of data silos while ensuring the privacy of data. Further advancement in the field of FL plays a vital role in different fields, especially healthcare. It shows that the quality of medical and clinical diagnosis significantly improved by adopting FL, which should be trained on medical data across multiple hospitals. FL is also a type of distributed ML that builds a collaborative model without compromising privacy. FL perfectly addresses the issue of privacy leakage that was faced before by the distributed ML. FL is a collaborative learning model in which multiple client machines train local ML models using their local data. The difference between distributed ML and FL are presented in Table 4.

The training process of FL is simple. First, the central-server broadcast default gradient and a generic global model for initial training. Each collaborator client downloads the gradient information and generic model and modifies all the information. The global model is trained on local data, and then gradient information is uploaded to the central server using encryption techniques. The central server performs aggregation of all the uploaded gradients and updates the global model.

ML has significantly improved the efficiency of the medical industry, especially in the decision-making process and clinical trials. Training an ML model always has a privacy risk regarding the model itself and the associated data. The problem can be resolved by federated learning (FL). In the case of an encryption system, the file is first encrypted by a specific algorithm and then transferred. But in the case of FL, only the parameters are transferred instead of sharing medical data. The localized model training and transferring of the parameters can be encrypted by different algorithms like Homomorphic Encryption (HE), Differential Privacy (DP), and Secure Multi-Party Computation (SMPC) [65].

In FL it is not essential to upload data for model training. FL just focuses on model training but without considering data samples. FL is an ideal framework that develops an ML model that ensures sensitive medical records [66]. Gao et al. [67] proposed SVeriFL model, the model was based on multi-party security and BLS signature. The participants uploaded the parameters to the server and the server testified the aggregation results.

Currently, there are two major and famous frameworks of FL, WeBank Fate and TensorFlow framework. Besides there are some other frameworks like Flower [68], which is a relatively new framework that isolates itself from others based on new facilities regarding large-scale FL experiments. FL is considered to be the next-level privacy-preserving framework for AI applications.

FL has addressed a lot of issues in the healthcare domain, like the lack of publicly available datasets and some other issues like privacy and confidentiality. Mostly horizontal FL and federated transfer learning have been used to train the model, which achieved the best results on textual medical

**TABLE 4. Difference between the Distributed ML and FL.**

| Framework | Type | Features | Benefits | Limitations |
|---|---|---|---|---|
| Distributed ML | Parallelism model | Same type of model is split into different steps and trained on the different client machines | This technique increases the computational speed | In case of network fluctuation, data packets may be lost leads to experimental delay |
| Distributed ML | Data Parallelism | The same model is deployed on different client machines with data and finally, training outcomes from all clients are aggregated | Results in decreasing model's run-time | Assigning different data to each client leads to maximum error |
| Distributed ML | Hybrid Parallelism | Switch between different GPUs and performing data-parallelism to train model on client machines | The approach reduces model's run-time and increases the computational speed | Excessive memory and GPUs consumptions |
| FL | Horizontal FL | More overlapping of data-features and less overlapping of data-samples | Increases model's training on different samples of data | Model's accuracy may be compromised because of single database |
| FL | Vertical FL | Less overlapping of data-features and more overlapping of data-samples | Increases model's training on different features of data | Model's accuracy maybe compromised because of single database |
| FL | Federated transfer-learning | Less overlapping of data-features and less overlapping of data-samples | Trains the model with different data-samples and features | Model's accuracy may be compromised because of single database |
| BC based FL | BC FL | Introducing BC in FL | Improves the accuracy and the security of the system | Take a long time and also needs hardware support |

datasets. Most researchers have combined BC technology with FL and achieved effective outcomes. Researchers have combined FL with ML and DL and then applied it to medical and clinical diagnosis. FL have a lot of practical applications like tumour detection [69], shadow segmentation [30], semantic segmentation [70] cancer prediction [71], EEG classification [72], grading prediction [73], genomic characterization [74], recurrence and quantification prediction, mood diagnosis [28], [75], [76], retinal-fundus imaging [77], melanoma detection, chest-X-rays analysis and COVID-19 diagnosis [9], [10], [78] and recently get a lot of attention. FL has a lot of advantages for medical applications, FL in healthcare needs the medical data of different diseases, it also depends on the numbers of FL clients and the selection of a federated-aggregation algorithm. Mostly the selection of a federated aggregation algorithm has a direct effect on final outcomes. The number of clients affects the accuracy and the time of the model's training.

The initial training of the FL model is unable to provide good performance, but after some iterations, the performance starts improving [21]. Gao et al. [72] have used FL for the detection of EEG stimulus maps using MindBigData. They used the dataset of 2-3 devices and achieved better convergence speed and accuracy. Nguyen et al. [10] proposed the FedGAN algorithm and tested the algorithm for the diagnosis of COVID-19. The algorithm has performed well on two datasets and saves time for model training and running time. Lee et al. [21] have used three datasets ECG, MIMIC-III, and MNIST and predict patient mortality rates. the experimental results have demonstrated FL performs well by increasing the number of training rounds. The results show FL is a suitable framework to apply to healthcare datasets to solve real-world problems. Sheller et al. [79] validated FL by using MRI and BraTS 2017 datasets, after training a quality model it achieved 99% accuracy on 10 rounds.

Dou et al. [24] used CT images to detect positive cases of COVID-19 by using CNN and federated transfer-learning algorithms. The study shows FL model consistently performed well in all the metrics. Huang et al. [66] have used the CBFL model to predict patient mortality and survival chances, the mortality prediction has been estimated by training and testing patients' drug characteristics from the same hospital. The results show centralized learning performed well, but the CBFL outperformed when tested on 10 different hospitals. Peng et al. [28] gathered data by micro-blogging posts and questionnaires trained the CAFed algorithm, and split the training data across 10 different devices to improve the convergence rate and overcome the communication cost. Yang et al. [80] gathered COVID-19 patient data from Japan, Italy, and China split the data among two clients, and performed joint training the results achieved good accuracy.

**TABLE 5.** Research studies based on DL and FL in healthcare.

| Article | Dataset | Healthcare Application | Model | Outcomes |
|---|---|---|---|---|
| A. Vaid et al., [84] | EHR | Predict the mortality rate of COVID patients | Used FL and MLP model | The aggregation perform well than the centralized model |
| I. Dayan et al., [85] | EMRs data | Oxygen therapy | EXAM | By adopting the federated model have improved 34 % accuracy |
| A. Sadilek et al., [86] | SARS-CoV2 and MIMIC-III | Prediction of survival rate | FL and DNN model | Achieve higher and effective performance |
| Y. Huang et al., [87] | MRI | Classification of Alzheimer disease | FedCM | Achieve good results on dichotomous method |
| H. Lee et al., [88] | Ultrasound image dataset | Prediction of disease | ResNet-50 and FL | Both model perform equal results |
| M. Lu et al., [89] | Ultrasound image dataset | Prediction of breast cancer | FL | Perform best |
| H. Elayan et al., [8] | Dermatology dataset | Detection of skin diseases | DFL | Achieve good accuracy and AUC |
| B. Han et al., [33] | Dermatology dataset | Detection of skin diseases | DFL | Better results in terms of robustness |

**TABLE 6.** Questions for initial quality review.

| S. No | Questions | Response |
|---|---|---|
| IQ1 | Is the publication related to BC, DL or FL in healthcare | Yes / No |
| IQ2 | Does the research proposes a new concept, proposal, framework, or model or modify the existing model, implement a prototype or perform a novel objective-based study | Yes / No |
| IQ3 | Is the study have published in a scholarly journal, book section, white paper or conference | Yes / No |
| IQ4 | Is the study provide the basic level of details that are essential to answer the research question | Yes / No |

Feki et al. [81] gathered COVID-19 image-based data and divided data into 80-20 ratios for training and testing and divided the data among four clients for simulation purposes. The technique provided effective outcomes. Borger et al. [75] collected data from four care-wards of two healthcare institutions and performed a training simulation to validate the mental condition of psychiatric patients. This was the very first study to use NLP and FL on clinical textual data. Bercea et al. [82] proposed an FL-based unsupervised model named FedDis and evaluated its performance for the identification of abnormal brains using MRI images. Initially, 1532 images were used to train the model and testing was performed across five institutions by which the model improved to 99.74% accurate results for abnormal segmentation, but on local training model just scored 40.45% accuracy. The FedDis model is beneficial for sharing abnormal health data across the institute and it improves the abnormality detection rate to 227% for trained tumour detection. It improved the detection rate to 77%. Ahmed et al. [83] have proposed an emotional lexicon and structural hyper-graph for the representation of words. They have adopted FL based model for the detection of mental health symptoms.

## III. REVIEW METHODOLOGY

We have adopted the Systematic Literature Review (SLR) guidelines [90] and Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) methodology [91] and conducted this research study. SLR refers to the methodology that discovers, analyzes and accesses recent research literature related to the subject field. The research papers were searched from academic repositories like IEEE, WOS, Scopus and PubMed in March 2023, by the keywords of ''blockchain in healthcare'', ''Deep Learning in healthcare'', ''federated learning in healthcare'', ''securing patients records'', ''digital patient records'', and ''electronic health records''. By using all these keywords we designed search string the complete details discussed in the subsection III-B, by using the search string we get different results from different databases. The designed search string provides 199 research articles in total. The next step was to filter out the articles and we applied initial quality review that was based on four questions as discussed in Table 6. The complete details of initial quality review have discussed in subsection III-C. Continuing to the filtering process, we define the quality assessment of articles, the details are discussed in subsection III-D, based on some quality assessment questions that are discussed in Table 7. And the results of research articles that answers the Quality Assessment questions have mentioned in Table 8. And then we perform the final evaluation and selection of articles based on inclusion and exclusion policy, that are discussed in subsection III-E. The final results of filtered articles have discussed in subsection III-F.

### A. DEFINING THE RESEARCH QUESTIONS (RQ)

First, we define the RQs for SLR that answer the following RQs, what architectural mechanism is used to ensure

**TABLE 7.** Questions for quality assessment.

| S. No | Description |
|---|---|
| QA1 | Is the research articles according to our domain and fulfil the purpose of our study? |
| QA2 | Is the research study proposed a prototype, model or framework achieve the security and confidentiality of healthcare data? |
| QA3 | Is the research article have a main contribution with respect to our research aims? |
| QA4 | Do the findings of the research article have capable of making a comparison with other existing techniques? |
| QA5 | Does the model proposed in the article have synthetic or real data that presented the output? |
| QA6 | Does the conclusion of the article is according to the title and objective of the research? |

interoperability and the security of EHMS using BlockChain technology?

**RQ1:** What are the architectural mechanism that has been used to support security and the interoperability of electronic health records using Blockchain technology?

**RQ2:** What are the different frameworks and tools that have been used to improve privacy and security generally in healthcare and specifically to secure the electronic health record using Blockchain technology?

**RQ3:** What are the open issues related to the usage of blockchain technology in the electronic healthcare domain and their future perspectives?

**RQ4:** What is the standardized way to store EHRs and handle big data in blockchain systems?

**RQ5:** How the FL preserve the privacy of EHRs in the healthcare domain?

### B. CONDUCTING THE SEARCH

The research questions were formulated after searching for a literature review in March 2023 from the year range 2018 to 2023, including the results being updated in March 2023. The search string includes TITLE-ABS-KEY((''blockchain in healthcare'' OR ''Deep Learning in healthcare'' OR ''federated learning in healthcare'') OR (securing AND patients AND records AND digital AND patient AND electronic health records)). Figure 1 shows the flow chart of our strategy. Additional tools have been used for the refinement of articles based on the context of the article, association with the interest of the study, and related search. And finally, we get the full text of all the research articles that are important for our SLR.

### C. INITIAL QUALITY METRICS

The Initial Quality (IQ) review process has based on the certain questions mentioned in Table 6, and questions IQ1, IQ2, IQ3 and IQ4 served as quality metrics for the initial review step. The title, abstract and full text of the articles have been scrutinized against the quality metrics.

### D. QUALITY ASSESSMENTS OF ARTICLES

Quality Assessments (QA) of selected articles are performed to obtain more specific research articles for SLR. The protocol introduced by A K-Peterson [39] is adopted for the process of QA and to discover the overall quality of selected articles. We formulated some questions discussed in Table 7. First, we read the full text of all the selected articles and calculated the quality score of each study on the basis of the QA question discussed in Table 7. The research studies having a quality score of more than 50 percent of the criteria are included in the final list. The QA1 and QA2 remain essential for all the research studies to qualify for the QA test. Based on independent QA scores, only 20 studies were selected as the primary study (because they strictly follow up the QA1 and QA2). Table 8 shows the QA results of the primary studies on the basis of the QA question discussed in Table 7. Each column (QS1 to QS6) in Table 8 shows whether the selected research study answers QA questions. If the selected study satisfies the QA question, it is granted a 1 (quality score). If it partially answers the QA questions, it is marked a 0.5 (quality score). If it fails to answer the QA questions, it is allocated a 0 (quality score). Overall 46 research studies have been selected including the primary studies that achieved more than 50 percent of the quality score.

### E. EVALUATION AND THE SELECTION OF RESEARCH ARTICLES

We define Inclusion (I) and Exclusion (E) for adding/removing the research articles that are relevant/irrelevant to SLR.

#### 1) INCLUSION CRITERIA

We have performed the inclusion of articles based on the following supporting arguments.

**I1:** The research articles must have been published in the last five years.

**I2:** If it discovers the research articles that show the same research study, then the most recent research article will be selected.

**I3:** If a research article shows multiple studies, each research study will be evaluated individually.

**I4:** If a research study has full and short versions, we prefer to keep the full version.

**I5:** Peer-reviewed, proceeding conferences, surveys, and research articles are included.

#### 2) EXCLUSION CRITERIA

**E1:** Papers not written in the English language.

**E2:** Technical reports, abstracts and book chapters.

**TABLE 8.** Overall result of the quality assessment.

| Ref. Studies | QS1 | QS2 | QS3 | QS4 | QS5 | QS6 | Quality Score |
|---|---|---|---|---|---|---|---|
| [64] | 0.5 | 0 | 0.5 | 1 | 1 | 1 | 4 |
| [36] | 1 | 1 | 1 | 1 | 1 | 1 | 5 |
| [37] | 1 | 1 | 1 | 1 | 0.5 | 1 | 5.5 |
| [38] | 1 | 1 | 0.5 | 0.5 | 0.5 | 1 | 4.5 |
| [92] | 1 | 0.5 | 0.5 | 0.5 | 1 | 1 | 4.5 |
| [40] | 1 | 0.5 | 1 | 1 | 0.5 | 0.5 | 4.5 |
| [26] | 0.5 | 1 | 1 | 0.5 | 1 | 1 | 5 |
| [27] | 1 | 0.5 | 1 | 0.5 | 1 | 0.5 | 4.5 |
| [34] | 1 | 1 | 0.5 | 1 | 0.5 | 1 | 5 |
| [35] | 0.5 | 0.5 | 1 | 1 | 1 | 0.5 | 4.5 |
| [29] | 1 | 1 | 0.5 | 0.5 | 1 | 1 | 5 |
| [93] | 1 | 0.5 | 1 | 1 | 0.5 | 0.5 | 5.5 |
| [94] | 0.5 | 1 | 0.5 | 1 | 1 | 1 | 5 |
| [25] | 1 | 0.5 | 1 | 1 | 0.5 | 0.5 | 5.5 |
| [11] | 1 | 1 | 0.5 | 0.5 | 1 | 1 | 5 |
| [11] | 1 | 1 | 0.5 | 1 | 1 | 0.5 | 5 |
| [95] | 0.5 | 0.5 | 1 | 0.5 | 1 | 1 | 5.5 |
| [96] | 1 | 1 | 0.5 | 1 | 1 | 0.5 | 5 |
| [97] | 0.5 | 0.5 | 1 | 0.5 | 1 | 1 | 5.5 |
| [98] | 1 | 1 | 0.5 | 1 | 0.5 | 1 | 5.5 |
| [99] | 0.5 | 1 | 1 | 0.5 | 1 | 1 | 5 |
| [100] | 1 | 1 | 0 | 1 | 1 | 0.5 | 4.5 |
| [101] | 0 | 1 | 0.5 | 1 | 1 | 1 | 4.5 |
| [102] | 1 | 0.5 | 1 | 0.5 | 1 | 0.5 | 4.5 |
| [103] | 1 | 1 | 0.5 | 1 | 1 | 1 | 5.5 |
| [104] | 1 | 1 | 0.5 | 1 | 0.5 | 1 | 5 |
| [105] | 0.5 | 0.5 | 1 | 1 | 1 | 1 | 5 |
| [99] | 1 | 1 | 0.5 | 1 | 0.5 | 1 | 5 |
| [106] | 0.5 | 1 | 1 | 0 | 1 | 1 | 4.5 |
| [52] | 1 | 0.5 | 1 | 1 | 0.5 | 0.5 | 4.5 |
| [107] | 1 | 0 | 1 | 0.5 | 1 | 1 | 4.5 |
| [108] | 0.5 | 1 | 0.5 | 1 | 0.5 | 1 | 4.5 |
| [12] | 1 | 1 | 1 | 0.5 | 1 | 1 | 5.5 |
| [51] | 1 | 0.5 | 0.5 | 1 | 0.5 | 1 | 4.5 |
| [18] | 0.5 | 1 | 1 | 0 | 1 | 0.5 | 4 |
| [109] | 1 | 0 | 1 | 1 | 0.5 | 1 | 4.5 |
| [14] | 1 | 1 | 0.5 | 0.5 | 1 | 1 | 5 |
| [110] | 1 | 1 | 1 | 1 | 0.5 | 1 | 5.5 |
| [111] | 1 | 1 | 1 | 0.5 | 1 | 0.5 | 5 |
| [112] | 0.5 | 0.5 | 1 | 1 | 0.5 | 1 | 5.5 |
| [113] | 1 | 1 | 1 | 1 | 1 | 0.5 | 5.5 |
| [93] | 0.5 | 1 | 0.5 | 1 | 1 | 1 | 5 |
| [114] | 1 | 1 | 1 | 1 | 1 | 0.5 | 5.5 |
| [115] | 1 | 1 | 0.5 | 0.5 | 1 | 1 | 5 |
| [116] | 0.5 | 0.5 | 1 | 1 | 0.5 | 1 | 5.5 |
| [117] | 1 | 0.5 | 1 | 1 | 1 | 0.5 | 5 |
| [118] | 0.5 | 1 | 1 | 0.5 | 0.5 | 1 | 5.5 |

**E3:** During the first phase of reading titles and abstracts for the papers, we excluded healthcare research articles that do not have a general focus on BC and Artificial Intelligence.

**E4:** During the second phase of reading the full articles, we excluded articles that do not show technical aspects of BC and Artificial Intelligence in Healthcare.

**E5:** We do not focus on papers not related to health, architectural mechanism, interoperability and security.

### F. ANALYSIS AND THE PRESENTATION OF RESULTS

The chief aim of conducting SLR is to explore the mechanisms that are used to improve the interoperability and the security of EHRs using BC and DL. Figure 1. describes the refinement process. Initially, when we searched the designed string on selected databases, we obtained 199 research articles. It was discovered that 18 articles are duplicated in multiple databases. The duplicate articles were eliminated, and we added one copy of each article in the final record. 5 articles have been removed for other reasons (out of scope or paper not in English). In the next step, 176 articles remained and were further analyzed. Next, inclusion-exclusion was applied to 176 articles. We read the title and abstracts of 176 articles and selected the articles that aligned with and satisfied our basic requirements. Out of 176 articles, 68 were excluded after reading the title and abstract. Next, 108 articles remained, and after reading the full text of the articles, we again applied the inclusion-exclusion criteria. Out of 108, 62 articles were again eliminated as they did not show the technical aspects of our requirements. Finally, 46 articles were used as evidence to answer the designed RQs.

## IV. ANSWERS TO THE RESEARCH QUESTIONS

Taking into consideration **RQ1: What are the architectural mechanism that has been used to support security and the interoperability of EHRs using BC technology?** Based

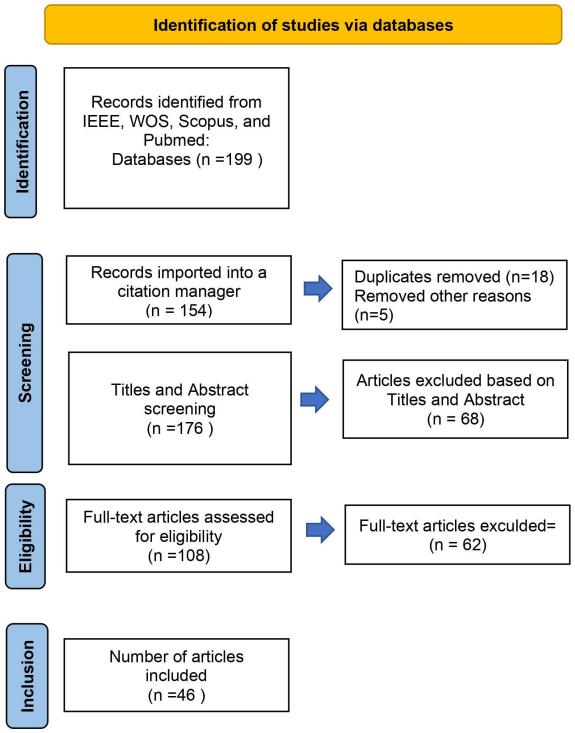



**FIGURE 1.** PRISMA methodology.

on the search string, we focused on the research articles that addressed security and interoperability in healthcare, and the studies proposed a BC-based architectural mechanism. 5 papers used a BC-based framework for their proposed solution. Amongst these, 4 solutions were based on meta-models, and the Intermediary framework was used in 3 research articles. We focused on answering RQ1 based on the novelty and rapid development of architectural mechanisms in the research area, and their impact on our solution is presented in section IV-A.

Modeling architecture uses a high level of abstraction known as views. The frameworks commonly used different viewpoints to create a specific view that describes the elements of the system's model. Specific frameworks were adopted for different natures of research studies. Some studies focused on enterprise architecture, decentralized frameworks or distributed systems [119]. For example, [93] proposed a BC-based cross-domain framework to distribute radiological images. Reference [114] discussed a BC-based framework to distribute the data of oncology patients. Reference [115] proposed an EHRs framework based on the BC network, which stored and distributed the EHRs data and also maintained a single version of the truth. Stakeholders request to access patients' data based on the Smart Contract (SC). On approval, they can make a transaction to the distributed ledger. Commonly user interfaces provide ease to users to interact with software. People and other software applications use the software. This interface is called an Application Programming Interface (API) [120]. APIs are the best and simple way to connect, extend or integrate a module in a software system. In most cases, the APIs are publicly available for researchers and developers. They are a robust link for off-chain and on-chain data management [115]. The methodology based on the smart contract was proposed in [121], which provided open-source access to the developers to regenerate the structural code and modify the smart contracts according to their interests. They also proposed feature and domain-oriented developmental methodology to generate and analyze Domain Specific Language (DSL) and deploy SC among BC networks. They performed three case studies and validated the files created by DSL. In the same study, they described that Model-Driven Engineering (MDE) is not a suitable tool for creating and managing the overall process of generating DSL.

The conclusion of the above-mentioned discussion is as follows. 11 research articles used the BC framework and suggested the solution to the assigned task presented in Table 9. According to the study [122], the evaluation of software reuse was started by implementing source code from scratch. A framework can be generated for a specific domain. Before designing, we should gain enough domain knowledge. The next step in the evaluation process is DSL. Once we have a framework, the code can be generated by DSLs that improve productivity and quality, interoperability and maintenance of the system. **RQ2: What are the different frameworks and tools that were used to improve privacy and security generally in healthcare and specifically to secure the EHRs using BC technology?**

In Table 9 and Table 10, we have discussed different frameworks of BC that were used in the existing research in the healthcare domain. In Table 9, we discussed different architectural mechanisms like meta-models, frameworks and intermediaries used in BC-based healthcare systems. We also mentioned the consensus algorithm, BC type, BC platform and the state of the proposed research study. Either it was a proposal, prototype, case study, or experimental study. In Table 10, we have discussed research articles with respect to their research direction. The technique used, such as the framework, and the dataset used to perform experiments and research achievements. All these BC-based tools and frameworks have various features. The salient feature is the smart contract that protects data tampering from unauthorized access, as discussed in Table 10. Research studies adopted smart tools and frameworks to protect and secure data access, ensuring that patients own their information, avoid any unauthorized third party from accessing confidential data, protect the gathered data by smart wearable devices and sensors, securely distribute the data among stockholders, usage of AI and ML for disease prediction and diagnosis, advanced encryption schemes to encrypt data, secure management of EHRs data, etc.

**RQ3: What are the open issues related to the usage of BC technology in the Electronic Healthcare Domain and their future perspectives?**

The open issues in the existing BC-based healthcare systems need to be addressed. In Table 11, we have identified and summarized open issues in existing systems. There are

**TABLE 9.** Different aspects of BC framework for various research studies.

| Paper | Architectural mechanism | | | Blockchain type | | | | Platforms | | | Article state | | | | Consensus Mechchanism |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Intermediary | Meta Model | Framework | Public | Private | Consortium | Authorized | Ethereum | Hyperledger | Other | Proposal | Experimental | Prototype | Case Study | |
| A. Rehman et al., [36] | | | ✓ | | ✓ | | | | | ✓ | | ✓ | | | PoW |
| X. Yang et al., [37] | | ✓ | | | | | ✓ | ✓ | | | | ✓ | | | Not mentioned |
| D. Hoang et al., [38] | | ✓ | | | | ✓ | | | ✓ | | | | ✓ | | CFT |
| G. Hegde et al., [92] | | | ✓ | ✓ | | | | ✓ | | | | | ✓ | | Not mentioned |
| V. Marichamy et al., [40] | | | ✓ | ✓ | | | | | | ✓ | | ✓ | | | PoW |
| A. Razzaq et al., [26] | | ✓ | | | ✓ | | | ✓ | | | | | ✓ | | Not mentioned |
| D. Shah et al., [27] | | | ✓ | ✓ | | | | ✓ | | | | | ✓ | | Not mentioned |
| G. Al-Sumaidaee et al., [34] | ✓ | | | | ✓ | | | | ✓ | | | | | ✓ | Not mentioned |
| M. Abouali et al., [35] | | ✓ | | ✓ | | | | ✓ | | | | | ✓ | | Not mentioned |
| J. Jayabalan et al., [29] | | | ✓ | ✓ | | | | ✓ | | | | | ✓ | | Not mentioned |
| Y. Bae et al., [94] | ✓ | | | ✓ | | | | | | ✓ | | | | ✓ | DPoS |

**TABLE 10.** Research studies that used different frameworks and tools to improve privacy and security in healthcare system.

| Article | Direction | Technique | Framework | Dataset | Achievement |
|---|---|---|---|---|---|
| A. Rehman et al., [36] | Secure Healthcare System | IDS, FL and BC-based model | BC and FL based real-time extreme-learning system (RTS-DELM) | Parkinson's disease & NSL-KDD dataset | Proposed RTS-DELM scored 93.22% accuracy for disease prediction and 96.18% accuracy for intrusion detection |
| X. Yang et al., [37] | Signcryption of EHRs using BC and Cloud server | Combine BC and attribute-signcryption-scheme | BC based verifiable and outsourced-attribute based Signcryption-scheme (BVOABSC) | Not mentioned | Performance evaluation and the security analysis showed the proposed scheme achieved high efficiency and is more secure |
| H. Hoang et al., [38] | Securing personal-health records | Proposed HoloCare system based on BC using hyperledger fabrics with CFT consensus algo | Holocare system based on Hyperledger fabric based | Not mentioned | The performance evaluation using hyperledger caliper-tool and made 10000 invoke transaction with min latency of 0.06s |
| G. Hegde et al., [92] | Secure storage of EHR using BC and disease prediction using ML | Used three off-chain storage methods IPFS, Storj and CosmosDB for EHRs-data storage and XGBoost for disease prediction | Framework based on IPFS for data storage | Chronic Kidney Disease | The average time for IPFS data storage is 0.11s and 0.095s for data access |
| J. Jayabalan et al., [29] | IPFS Off-chain BC for ensuring Privacy and Security of PHRs | Used AES-128 and encrypt PHRs and store on IPFS, also used multifactor authentication to avoid fake-node attacks | Framework based on IPFS for data storage | Not mentioned | Performance evaluation of proposed framework with SOTA work on the basis of storage-retrieval time-analysis, the ratio of storage-compression and exhibited good results |
| T Benil et al., [25] | Enhancing security of e-records in cloud system using BC | Used modified ECC for data-encryption | Adopt random oracle model | Not mentioned | Performance evaluation based on auditing time, computation-cost, verification delay and showed good results |
| V. Marichamy et al., [40] | Securing Medical Records using BC | Used cryptographic hash-generator (CHG) to secure user-key, for the encryption of medical-data applied discrete-shearlet transform (DST). Data was stored in BC. Adopted grey-wolf optimization algorithm to find optimum request | BC-based framework named BC-SMR-BD-GA-DWT | Not mentioned | The proposed system was executed in Hadoop platform, the proposed model scored 99.08% accuracy on 10Mb of data, and showed 12.5% of lower-attack level using 10Mb of data |

some common issues related to scalability and privacy of EHRs in most of the research studies.

**RQ4: What is the standardized way to store EHRs and handle big data in the BC systems?**

The format of data and interoperability standards always remained an issue for storing, accessing, and distributing EHRs. Most research works showed that researchers did not follow the standards and guidelines of Health Level Seven (HL-7) and Fast-Health Interoperability Resources (FHIR). Some researchers applied standard principles in their proposed research solution. Most of the researchers considered HL-7 and FHIR while designing the EHRs data format [116], [117], [123], [124]. Some researchers were bound to the guidelines of the Health Insurance Portability and Accountability Act (HIPPA) [118], [125], [126], [127]. Very few researchers follow the principle of HL7 [128],

[129]. In the Indian health sector, the standard introduced by the National-eHealth Authority (NeHA) was practiced as a regulatory and promotional standard for an organization and applied in [18].

The Virtual Medical Record (VMR) used in [110] is a standardized and simplified data model specially designed for clinical decisions. A research study [97], [130] showed the researcher only described some standards like HIPPA, ISO 18308: 2011 and HL-7, but the study was far behind from actual implementation of these standards. In the research [131], the author implemented the principle of OpenEHR. No other research work used this standard. For big data in BC, there are a lot of hospitals, clinical centers, and patients that continuously generate big data. People take medical care worldwide, and EHRs are generated for diagnosis purposes. It is a challenging task to handle such an enormous amount of data. For BC, handling big data is a challenging task since it is much more expensive to store such huge data over the BC. Initially, BC technology was developed to store a small amount of information related to financial transactions. Later on, many researchers came up with different ideas on how to enhance the data storage capability of BC. Most researchers did not consider the scalability issues of the BC technology while storing the data. Usually, big data is stored over cloud storage or local databases, and this local storage address is linked to BC.

In the reviewed research articles, more than 60% of the research articles did not consider the storage issues of big data. Only the articles [11], [52], [99], [105], [106] encountered the storage issues of big data in their studies without mentioning the storage services. The studies [12], [14], [18], [51], [92], [107], [108], [109] used IPFS for data storage and linked the storage address to the BC. In [110], [111], and [112], researchers have adopted Amazon Cloud Services for data storage and linked the storage address to BC. Very few studies used local databases to store EHRs data and link to the BC. To handle the scalability problems, some studies also used off-chain storage. The solutions to handle big data issues are significant, but it will need a lot of work to achieve perfection.

In light of all aforementioned references, we conclude that BC-based EHRs systems still have several issues related to data formatting, authenticating, exchanging, storing, uploading and handling big data. It is due to the evolutionary nature of BC or the lack of a standard platform or methods. BC is a promising platform for EHRs data management, but it may take a long time to achieve a stable rank as a standardized framework that solves the aforementioned issues.

**RQ5: How the FL preserves the privacy of EHRs in the healthcare domain.** During the training of the DL model, it is necessary to ensure security, privacy,and additional measures to extend and strengthen the privacy protection mechanism when the attack is launched. FL is a framework that secures the process of model training. In FL, secure aggregation is the most prominent topic and more significant in solving the issue of an untrusted central server. Secure aggregation is also helpful in preserving the privacy of EHRs. There are different ways to perform secure aggregation, like multi-party computation, homomorphic encryption, differential privacy, etc., as discussed in Table 12. The algorithm focuses on privacy preservation in FL and uses these three methods for secure aggregation.

SMPC is the cryptographic technique that serves as an underlying framework for the privacy protection of FL. The cryptographic technique plays a major role in securing the data, ensuring integrity, privacy and data authentication. Figure 2 shows the SMPC of the FL framework.

SMPC accurately performs the federated computation while preserving privacy and adopts the distributed nature of FL. It calculates collaborative data by multiple clients in a secure and distributed manner [127], [136], [137]. SMPC enables segmentation between the shared encrypted data so that no participating clients can retrieve all the data. After calculating the result, the clients are unable to see the data. Data is recovered by a consensus process. SMPC also consider dishonest behaviour by illegitimate clients. Attackers design an attack to steal data through illicit clients or changes in computing tasks. SMPC protects the privacy of FL in the medical domain. The second method for secure aggregation is Differential Privacy (DP). It provides data security and makes sure all sensitive healthcare-related data is secure. DP protects the data from all adversaries who want to target the confidentiality and privacy of data. When DP is applied with FL, it secures the leakage of weights. DP adds partial noise to avoid the computationally overpowered adversaries. DP method is best for large datasets rather than small datasets. Because during the training process, if partial noise is added to a small dataset, it highly affects the results.

The noise padding in the dataset is important. It preserves the global distribution of data. Moreover, it reduces the probability of individual identification with the information. If the adversaries cannot infer whether a specific individual is part of the dataset, it shows the dataset is private to certain degrees. DP is useful for avoiding the re-identification attack, adding partial noise to the dataset and protecting sensitive data. In FL and the healthcare domain, DP is the most commonly used privacy algorithm. It is used to develop a collaborative model to train medical images or textual data. Lu et al. [89] utilized DP with attentional multi-learning and secured the computational pathological data. The study shows multi pose-learning develops a DL model that effectively distributes the data by avoiding the direct sharing of data and certain complexities. Chang et al. [138] used an adaptive DP algorithm for data and privacy protection and adopted a gradient-verification consensus protocol for poison-attack detection. The study was applied to protect the data of diabetic patients. Li et al. [139] achieved privacy protection by using three-layered architecture, used IoT devices with DP mechanisms and FL security mechanisms

**TABLE 11.** Open issues identified in recent research studies regarding scalability and privacy.

| Open issues | References |
|---|---|
| Scalability issues | [37] [92] [40] [25] [27] [35] [29] [11] [132] |
| Privacy issues | [36] [11] [34] [95] [96] [97] [98] [99] [100] |
| Lack of standard evaluation method | [35] [29] [11] [101] [133] |
| Lack of standardised performance metrics | [11] [27] [34] [35] [29] |
| Lack of decentralised storage | [37] [38], [134] [92] [11] [25] [27] [34] [35] [29] [102] |
| Lack of decentralised access to data | [36] [38] [11] [25] [27] [29] |
| High computation cost | [40] [11] [135] [97] [100] |
| Lack of standardised authentication methods | [38] [11] [34] [35] [103] [104] |
| Lack of decentralised consensus | [37] [92] [40] [25] [27] [104] |
| Studies without smart contract | [36] [37] [40] [38] [11] [25] [35] [95] [101] [102] [96] [104] |

**TABLE 12.** ML and DL techniques for detection of IoMT attacks.

| Algorithm | Benefits | Drawbacks |
|---|---|---|
| Secure Multi Party Computation (SMPC) | Provide a fast training model | It is vulnerable to the attacks and hardware support is required for SMPC. |
| Differential privacy | It keeps the data secure | Because of noise, model performance might be affected |
| Homomorphic encryption | Encrypts the data and provides high security to data | Because of the high computational cost, the efficiency might be low. |

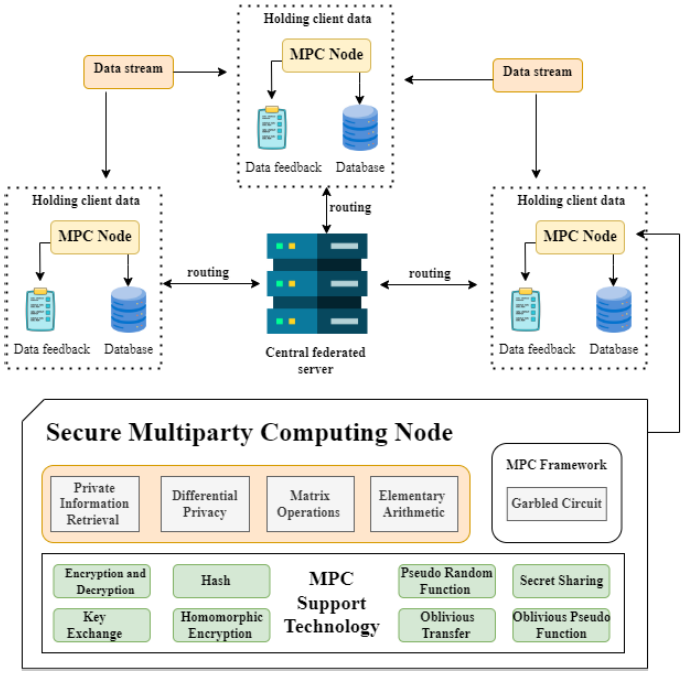



**FIGURE 2.** Secure multi party-computation.

and applied this three-layered architecture for Alzheimer's disease detection.

Wu et al. [25] proposed a research study by adding artificial partial noise into the local client's dataset to preserve the user's privacy and solve a multi-dimensional problem. Malekzadeh et al. [26] proposed a study in which DP homomorphic encryption and stochastic gradient-decent (SGD) were used for secure aggregation, distribution of medical data and privacy preservation. DP has a significant role in protecting the cost of partial loss of precision. It is a powerful technique that guarantees privacy preservation for FL in the medical domain. The third method usually used for secure aggregation in FL is Homomorphic Encryption (HE), as presented in Figure 3 and it is different from conventional encryption techniques because it performs certain operations on original data and encrypted data without requiring decryption [27]. In HE, the secret key was shared between the peers to secure the encrypted message. Only the peer nodes with the private key had a legal right to access the data. But nowadays, some cloud servers lose their sensitive information because they do not have any shared key, and the encrypted data is shared with third parties. Additionally, in some untrusted cloud servers, operators continuously try to identify the user's data until and unless the user ends the

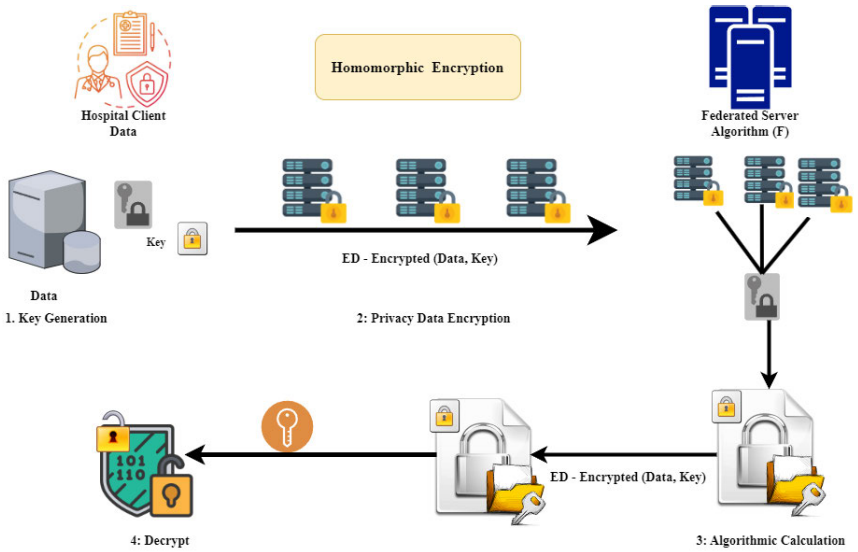



**FIGURE 3.** Homomorphic encryption.

relationship with the server [34]. Nowadays, HE is widely used in various frameworks of FL. Zhang et al. [35] proposed a Verifiable-Privacy-Preservation FL scheme (VPFL) for edge-computing. It posed low communication overhead and computational cost and attained high accuracy by preserving the gradient leakage in the transmission phase. Jia et al. [29] solved the problem associated with model reversal attacks and model inference-attack and proposed a scheme based on data protection and aggregation by using DP and HE for BC-backed FL in the IIoT systems. Two HE algorithms were used to solve the privacy issues of EMR data, lose some computational cost and achieve good security and privacy.

### *A. BC AND CLOUD-BASED FRAMEWORK TO ENSURE SECURITY AND INTEROPERABILITY OF EHRS*

Considering the research gap specifically targeting the security and interoperability of BC, we proposed a new framework based on smart contracts and cloud systems by using BC, aiming to secure EHRs. There are two advantages of using a cloud system to manage EHRs. First, it shares the patient records with other concerning clinical-centres easily. Secondly, it integrates all the essential groups of the clinical centres and helps medical staff to do their jobs efficiently [140], [141]. Besides the several benefits of cloud-system for managing EHRs, there are critical security threats, like the cloud system being hacked, data leakage, theft, and loss or exposure of the EHRs [123], [142]. For example, EHRs can be tampered with by an attack launched on the cloud server to hide the mal practising of medical staff or to benefit the insurance companies. Dishonest insurance organizations sometimes may hire an attacker to tamper or delete patient records to prove the pre-existence of health conditions. Medical malpractising or delay of diagnosis are also some of the reasons to claim medical insurance. Most of the time patients cannot claim malpractices because of the aforementioned reasons. Different research studies have proposed cryptographic methods to solve the security issues of EHRs [31], [32], [40], [113], [143], [144]. However, security issues still exist because of the centralized nature of cloud systems.

BC is a distributed ledger where data is shared, stored and exchanged across various stockholders. For e-health systems, medical data can be generated from different sources like hospitals, clinics and pathologists. In BC based EHRs management system, the patient data are stored in a distributed ledger (offered by BC). The process of storing data is known as a transaction. Before storing data in a distributed ledger, each transaction must be evaluated by miners. BC network rejects all unauthorized transaction that tries to tamper with the data. Therefore, no unauthorized person can access data stored in a distributed ledger. The chief element of BC is known as a smart contract, which is an executable program with a set of agreements. All the participants that are part of the network must adhere to the agreement, this way unauthorized third parties cannot store or access data in the BC network [145], [146], [147]. It is a challenging task to integrate BC into a traditional cloud system. First, we need to eliminate the central control of the data and should make decentralized access to data. It is only possible by integrating BC since data tempering is difficult in the BC network. The second issue is to adopt an appropriate BC network to store and manage EHRs. The third issue is that designing a system from scratch is quite laborious because it takes a long time, resources and effort. Therefore, we need to adopt the existing system as a backbone. As we have already discussed, integration of BC into a traditional cloud system is complicated. Therefore, we should adopt such a methodology that must be cost-effective, consumes fewer resources, deliver on time and not affect the current stockholders particularly. We design a methodology by adopting the basic framework

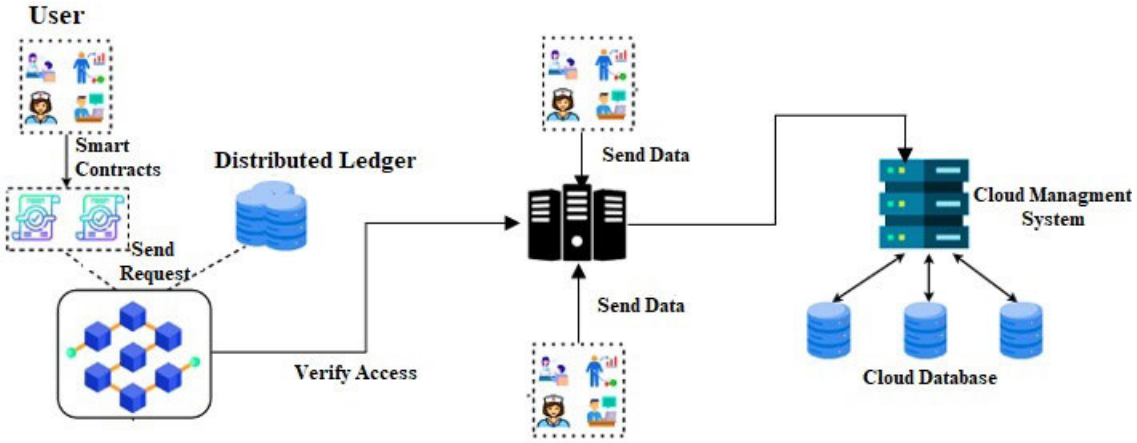



**FIGURE 4.** Proposed BC and cloud-based framework to ensure security and interoperability of EHRs.

proposed in [144] using a bottom-up approach that integrates BC into a cloud system using a public BC network. The proposed methodology ensures the interoperability and security of EHRs and efficiently manages and distributes EHRs using smart contracts. Figure 4 demonstrates our proposed framework. The doctors, nurses, pathologists and other clinical staff send the EHRs of the patients directly to the BC-wrapper (BC-w). The BC-w connects to the cloud-management layer, medical staff, users and public BC network. The job of BC-w is to generate, manage and validate the transactions. The medical staff sends an Initial Transaction (T1) to BC-w to communicate with the BC network. Then BC-w generates a transaction (TC) using a transaction generator (Tg). Another component of BC-w called transaction-validator (TV) sends Tc to the BC network for validation. The BC network validates the Tc using smart contracts and miners, and finally, the data is added to the BC network. The BC network sends a Validation Message (VM) to BC-w after the validation of TC. BC-w sends VM to the cloud for proceeding. Lastly, the EHRs data is stored on a cloud database.

## V. CHALLENGES AND TECHNICAL SOLUTIONS

The healthcare sector has undergone a change thanks to EHR, which offer better patient care, more efficient operations, and increased data accessibility. Over the past 20 years, there has been an exponential growth in the adoption of EHR systems, driven by government incentives and the desire for a more efficient healthcare ecosystem. But along with these developments come a number of security threats and flaws that security developers and healthcare professionals need to be aware of.

In this context, it is crucial to understand the multifaceted landscape of challenges and security vulnerabilities in current EHR systems. EHR systems present a number of challenges that need to be addressed in addition to their many advantages, which include improved care coordination, simpler access to patient data, and less paperwork, they also bring forth a range of issues that demand attention and mitigation.

- **Interoperability Dilemma:** The lack of interoperability among various EHR systems is also a problem. The healthcare delivery process becomes fragmented and inefficient when healthcare professionals use numerous EHR platforms for different areas of patient care and these systems are unable to connect with each other seamlessly. Trench form around patient data, which can lower care quality and result in expensive mistakes.
- **Data security and privacy:** Because EHR systems hold a wealth of private patient data, hackers find them to be appealing targets. Patients may be at risk of fraud, identity theft, and compromised medical record confidentiality due to security flaws in EHR systems. Additionally, violations may lead to legal ramifications and harm healthcare practitioners' reputations.
- **User authentication and access control:** It is crucial to make sure that only individuals with the proper authorization can access patient data. Inadequate user authentication and access control systems may result in data breaches, misuse of patient information, and unauthorized access. Finding the ideal balance between data security and healthcare professionals' ease of access continues to be a difficult task.
- **Human Error and Training:** Human error still occurs in the healthcare sector, and the implementation of EHR technologies has increased complexity. Insufficient training and low digital literacy among healthcare personnel can result in data input errors, inaccurate information sharing, and system abuse.

Besides all this, EHRs also have some other security vulnerabilities and severe security challenges, recently the U.S Department of Health and Human Services (HHS) has published a report concerning the recent cybersecurity risks for the EHRs. According to the report EHRs are still a top target for the cyber threat actors [148]. Noting that, the cyber threats, that mostly emphasizes on the security and

**TABLE 13. Open challenges and technical solutions.**

| Digital tool | Opportunities | Challenges | Technical solution |
|---|---|---|---|
| BC Technology | • Improve data access, data exchange and interoperability<br>• Ensure immutability and consensus<br>• Improve operation efficiency<br>• Improve the security and privacy of medical data<br>• Improve overall healthcare outcomes | • Low performance and poor scalability<br>• Expensive cost to maintain data-privacy<br>• Block-size and security vulnerabilities<br>• Number of nodes and protocol challenges<br>• Consensus and agreement are needed between participants of the network | • Decentralization of data and use of cryptographic techniques<br>• BC authorization, authentications and storage-optimization (VerSum, mini BC, Reference-pointer FHIRChain)<br>• BC modelling (OmniPHR, DeepLinQ, HealthChain, FHIRChain)<br>• Reading mechanism (catching system and short-term sharing of data)<br>• Writing mechanism (TrustChain, Fault-Tolerance protocol, Sharding, Tokenization, Practical-Byzantine, Cohort-algorithm, Smart contract) |
| Advance visualization | • Improve knowledge discovery<br>• Improve the information communication | • Larger EHRs dataset<br>• Diversity, temporal-complexity, evolving nature of EHRs data<br>• Lack of latest visualization techniques<br>• Low quality and completeness of data | • Lifelines<br>• VISITORS / KNAVE-II<br>• Methods implemented with the collaboration of other disciplines (engineering, computer science or genetics) |
| EHRs system in the austere-setting | • EHRs systems are required to deploy in austere-setting where paper-based storage and transports are not feasible<br>• Improve data completeness, quality and integrity<br>• Improve clinical management and diagnosis<br>• Provide a consistent and standardized practice environment<br>• Improve epidemiological analysis | • The settings could allow connectivity via expensive satellite connection, opportunistic internet connection or using the local-network<br>• Lack of interoperability<br>• For the austere setting, there are multiple EHRs systems but most of them are at the developmental stage | • OpenMRS has a good ability to integrate local EHRs systems with the MST medical-record<br>• The smaller EHRs system needs further developmental improvement in order to ensure interoperability (i.e. NotesFirst, QuickChart EMR, OpenMRS software, TEBOW, Project-Buendia, SmallList To-Go, iChart) |
| Clinical-Information Model (CIM) | • Provides structural and semantic interoperability of data b/w EHRs system | • Immutability of the latest modelling support-tools<br>• Different standards and models (i.e. HL7 V3, archetypes, OpenEHR, EN-ISO 13606) | • Openly shares the CIMs<br>• Collaboratively participates in the CIMs development<br>• A standardized methodology and best practice need to be developed for CIMs |
| De-identification tool, free text-processing or NLP | • Data-privacy preservation<br>• Enables the usage of EHRs for implementing personalized medicines, translational research, clinical, phenotype<br>• Beneficial for the unstructured data locked in the EHRs system<br>• Support the clinical management for effective outcomes | • De-identification put a negative impact while extracting the automated-information<br>• Errors, biases, poor data quality and involves privacy-issues<br>• The predominance of the rule-based over the ML-based NLP<br>• Involve difficult interpretability of ML methods and algorithmic biased<br>• Lack of standardized interoperable and poor generalizability | • ML-based technique based on conditional RF, DT, Maximum-entropy, SVM<br>• Needs to develop DL-based NLP methods for EHRs data-mining<br>• Share NLP models to GitHub to improve development and avoid duplication<br>• Development of ontology like biomedical and biological foundry |
| Deep learning | • Disease classification and detection<br>• Phenotyping and predicting clinical-events<br>• Data augmentation and data-privacy | • Irregularity and temporality of EHRs data<br>• Multimodal learning of EHRs is difficult because of the heterogeneous nature of data<br>• Difficult to find a standard way to label the EHRs data<br>• Lack of transparency and interpretability | • Gated architecture for extraction of temporal data<br>• Decomposition of LSTM to solve issues of time-irregularity<br>• Transfer learning and multi-task learning.<br>• Knowledge distillation and injection, attention mechanisms-based learning. |

privacy risks, hacking vulnerabilities, destroy or the loss of informative data, not limited to this, but there are some other top security threats like phishing attacks, frauds, malware and ransomware attacks, encryption blind spots, third party risks, insider or employee threats and the cloud threats. Recently a lot of breaches have been reported on the EHRs system. The BlackCat ransomware attacked on the NextGen EHR system, the NextGen system have more than 2500 healthcare provider customers, the attack had compromised hundreds of thousands patient files. According to another report by NextGen, that showed 1.05 million patients have affected by cyberattack [149]. Another EHR vendor EyeCare Leader reported that a security breach has affected 1.5 million patients. An Oregon healthcare system, Asante reported 8800 patients and their records have breached. The Health Insurance Portability and Accountability Act (HIPAA) recommended to use zero trust security model on the network and follow up the cybersecurity infrastructure agency (CISA) measures to protect EHR system against the critical cyber threats and finally to strengthens the organizational cyber posture.

The easiest way to avoid the cyber security vulnerabilities are to strictly follow up the HIPAA. Use the security checklists and rules, conduct the annual HIPAA risk analysis, regularly provide workforce cyber security training to prevent EHR breach. By following the guidelines can save the cost, time, and the loss of sensitive patient's data.

Table 13 describes some crucial digital tools like BC technology, that have a lot of opportunities regarding the data access, data exchange, interoperability, immutability, security and privacy of EHRs. But there are certain challenges like poor scalability, low performance, expensive maintenance cost, protocol challenges and block-sizes. The challenges can be solved in a technical way, by ensuring the decentralization of data, using cryptographic techniques, BC based authorization and authentication and storage optimization (VerSum, mini BC, Reference-pointer FHIRChain). The second digital tool is used for advanced visualization and it has opportunities to Improve knowledge discovery and information communication. There are also some challenges like Larger EHRs datasets, Diversity, temporal-complexity, evolving nature of EHRs data, Low quality of data. All these issues can be solved technically by the Methods implemented with the collaboration of other disciplines (engineering, computer science or genetics). The third digital tool is EHRs system in the austere-setting, and it has opportunities for the EHRs systems to deploy in austere-setting where paper-based storage and transports are not feasible, improve data completeness, ensure quality and integrity, improve clinical management and diagnosis. Also have some challenges like lack of interoperability, and the settings could allow connectivity via expensive satellite connection. The challenges can be solved technically by using OpenMRS that has a good ability to integrate local EHRs systems with the multiple subpial transaction (MST) for medical-record. There are some other digital tools like the Clinical-Information Model (CIM), De-identification tool, free text-processing or NLP and deep learning. All these digital tools have their own opportunities regarding the enhancement of EHRs security, management and privacy. The complete details of digital tools, their opportunities, challenges and the technical solutions are discussed in Table 13.

## VI. LIMITATION AND FUTURE DIRECTIONS

Healthcare systems based on blockchain and machine learning have a lot of potential, but they also have limitations, such as small sample sizes and a lack of real-world validation. Here are a few significant restrictions:

### A. SMALL SAMPLE SIZES

- **Data Scarcity:** Obtaining sufficient and diverse datasets for machine learning model training is one of the major issues in the healthcare industry. It might be challenging to develop reliable models in some healthcare applications when data is restricted to a small number of patients. This shortage may cause overfitting and result in an inaccurate representation of the larger population.
- **Bias and Generalization:** Models with low sample sizes may be biased and have poor generalization to various patient demographics, which could result in underprivileged patients receiving subpar care. Genetics, differences in healthcare procedures, and other variables that can influence results might not be taken into consideration by these models.

### B. LACK OF REAL-WORLD VALIDATION

- **Limited Clinical Validation:** Although machine learning models can yield good results on test datasets, there's a chance that their applicability to clinical practice will be limited. Validating models in real-world environments is necessary to guarantee their efficacy, safety, and generalizability.
- **Ethical and Legal Challenges:** Performing real-world validation frequently entails difficult ethical decisions as well as legal restrictions, especially when handling patient data that is sensitive. Adherence to privacy laws, including HIPAA, might provide difficulties.

### C. ISSUES WITH DATA QUALITY

- **Noisy and Incomplete Data:** Information related to healthcare is frequently erroneous, noisy, and incomplete. The effectiveness of machine learning models may be hampered by this. Misdiagnoses and projections might result from incomplete or inaccurate data.
- **Ethical and Legal Challenges:** Challenges with Data Labeling: Acquiring high-quality annotations can be challenging, and labeling healthcare data is a resource-intensive operation. Labeling mistakes have the potential to spread to machine learning algorithms.

### D. INTERPRETABILITY AND EXPLAINABILITY

- **Black-Box Models:** A lot of deep learning models, especially neural networks, are thought of as ''black-box'' models since it can be difficult to figure out why they predict certain things. Interpretability is essential to building trust and comprehending the decision-making process in the healthcare industry.

### E. REGULATORY AND COMPLIANCE BARRIERS

- **Regulatory Obstacles:** Adhering to regulations, such as the FDA's clearance of algorithms and medical devices, can be a costly and time-consuming procedure. Healthcare systems relying on machine learning may become less popular as a result.
- **Data Exchange and Privacy Issues:** Because of tight rules and privacy concerns, sharing highly sensitive healthcare data for research and validation purposes can be challenging.

### F. FAIRNESS AND BIAS

- **Model Bias:** Biases in the past healthcare data may be inherited by machine learning models that have been trained on it. These biases may have a detrimental impact on specific patient populations by producing unfair or discriminating results.

### G. INFRASTRUCTURE AND SCALABILITY

- **Scalability Issues:** As healthcare systems expand, it will become more difficult to scale blockchain and machine learning infrastructure to manage more users and data.

### H. INTENSIVENESS OF COST AND RESOURCES

- **Resource Requirements:** Creating and implementing machine learning models in the healthcare industry can be resource-intensive, involving large amounts of computer power, knowledgeable workers, and cash outlays.

A multifaceted strategy is needed to address these shortcomings, including better data collecting, managing biases, improving the interpretability of the models, and maintaining regulatory compliance. To assess the efficacy and safety of these technologies in healthcare, real-world validation via clinical trials and long-term investigations is necessary. Moreover, to fully utilize blockchain and machine learning in the healthcare industry while minimizing these drawbacks, cooperation between data scientists, regulatory agencies, and healthcare practitioners is essential.

One key avenue for future research is the development of integrated systems that seamlessly combine BC immutable ledger capabilities with machine learning's predictive analytics to proactively detect and prevent security breaches in EHRs. Additionally, still there is a need to explore the scalability and interoperability of BC solutions in healthcare settings to ensure widespread adoption. Furthermore, investigating the ethical and legal implications of deploying such technologies in healthcare, as well as addressing patient privacy concerns, will be pivotal in shaping the future landscape of EHRs security. Collaborative efforts among researchers, healthcare institutions, and technology providers will play a crucial role in realizing the full potential of BC and machine learning for enhanced EHRs security.

## VII. CONCLUSION

In this study, we have performed an SLR to identify the applications of BC, DL and FL in the healthcare domain. We adopted the PRISMA technique to systematically review the articles and extract meaningful information that provides more knowledge and answers our research questions. We focused on the architectural mechanism that solves the security and interoperability issues of EHRs using BC. The solution was proposed by mentioning the data extracted from SOTA literature that adopted BC architectural mechanisms to store EHRs. Secondly, the literature that made it possible to exchange electronic data between the stakeholder and avoid security breaches. We discussed different tools and frameworks that ensured the privacy and security of EHRs as researchers proposed several BC-based frameworks to improve the security and privacy of EHRs. In the light of selected articles, we mentioned different tools and frameworks commonly practised in the healthcare domain for security purposes. We described the open issues of BC technology in the healthcare domain and categorized open issues based on different factors. As each article has some flaws, but the overall factors that we discovered while performing the SLR were scalability and privacy issues, lack of standard evaluation and performance metrics, lack of decentralized storage and decentralized access to data, lack of standardized authentication methods, lack of decentralized consensus, High computation cost and the studies without smart contracts. We described the standardised way to store EHRs and handle big data in the BC systems. The identified issues can be addressed by integrating cloud-based storage in the BC system, i.e., by storing the actual data in the cloud and saving its hash in the ledger of BC. This way, data cannot be altered and will remain safe. Lastly, the issue of big data can also be resolved by integrating cloud storage in the BC system. In this SLR, we have proposed a BC and cloud-based framework that ensures the security and interoperability of EHRs. It handles big data and stores the EHRs in a standardised way. Finally, we discussed how the FL preserves the privacy of EHRs in the healthcare domain, and the solution has been proposed by different researchers, as FL provides a global model by aggregating the sum of all learned models. Rather than sharing patients' EHRs, a global model can be adopted that is already well-trained and used for testing purposes. In this way, the EHRs' privacy can be maintained.

## REFERENCES

[1] M. Boniol, T. Kunjumen, T. S. Nair, A. Siyam, J. Campbell, and K. Diallo, "The global health workforce stock and distribution in 2020 and 2030: A threat to equity and 'universal' health coverage," *BMJ Global Health*, vol. 7, no. 6, Jun. 2022, Art. no. e009316.

[2] V. Aggelidis and P. Chatzoglou, "Using a modified technology acceptance model in hospitals," *Int. J. Med. Informat.*, vol. 78, no. 2, pp. 115–126, Feb. 2009.
[3] A. Roehrs, C. A. da Costa, and R. da Rosa Righi, "OmniPHR: A distributed architecture model to integrate personal health records," *J. Biomed. Informat.*, vol. 71, pp. 70–81, 2017.
[4] N. Spence, N. Bhardwaj, and D. P. Paul III, "Ransomware in healthcare facilities: A harbinger of the future? *Perspect. Health Inf. Manage.*, pp. 1–22, Jul. 2018.
[5] M. Hassan, "A blockchain-based intelligent machine learning system for smart health care," Inst. Data Sci. Digit. Technol., Vilnius Univ., Vilnius, LT, USA, Tech. Rep., 2022.
[6] N. Thamer and R. Alubady, "A survey of ransomware attacks for healthcare systems: Risks, challenges, solutions and opportunity of research," in *Proc. 1st Babylon Int. Conf. Inf. Technol. Sci. (BICITS)*, Apr. 2021, pp. 210–216.
[7] T. Benson and T. Benson, "Using SNOMED and HL7 together," in *Principles of Health Interoperability HL7 and SNOMED*. Springer, 2012, pp. 267–280.
[8] H. Elayan, M. Aloqaily, and M. Guizani, "Sustainability of healthcare data analysis IoT-based systems using deep federated learning," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7338–7346, May 2022.
[9] R. Wang, J. Xu, Y. Ma, M. Talha, M. S. Al-Rakhami, and A. Ghoneim, "Auxiliary diagnosis of COVID-19 based on 5G-enabled federated learning," *IEEE Netw.*, vol. 35, no. 3, pp. 14–20, May 2021.
[10] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, and pA. Y. Zomaya, "Federated learning for COVID-19 detection with generative adversarial networks in edge cloud computing," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10257–10271, Jun. 2022.
[11] J. Vora, A. Nayyar, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, and pJ. J. P. C. Rodrigues, "BHEEM: A blockchain-based framework for securing electronic health records," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6.
[12] A. Al Mamun, M. U. F. Jahangir, S. Azam, M. S. Kaiser, and A. Karim, "A combined framework of interplanetary file system and blockchain to securely manage electronic medical records," in *Proc. Int. Conf. Trends Comput. Cognit. Eng. (TCCE)*. Cham, Switzerland: Springer, 2020, pp. 501–511.
[13] L. Coventry and D. Branley, "Cybersecurity in healthcare: A narrative review of trends, threats and ways forward," *Maturitas*, vol. 113, pp. 48–52, Jul. 2018.
[14] T. P. A. Rahoof and V. R. Deepthi, "HealthChain: A secure scalable health care data management system using blockchain," in *Proc. Int. Conf. Distrib. Comput. Internet Technol.*, Bhubaneswar, India, Cham, Switzerland: Springer, Jan. 2020, pp. 380–391.
[15] C. Burniske, E. Vaughn, J. Shelton, and A. Cahana, "How blockchain technology can enhance EHR operability," Gem| Ark Invest Res., New York, NY, USA, Tech. Rep., 2016.
[16] M. Hassan, J. Chen, T. Khan, U. Zukaib, and S. A. Mallah, "Future of edge computing for real time devices in Internet of Things," Wuhan Nat. Lab. Optoelectronics, Huazhong Univ. Sci. Technol., Wuhan, China, Tech. Rep., 2022.
[17] A. M.-H. Kuo, "Opportunities and challenges of cloud computing to improve health care services," *J. Med. Internet Res.*, vol. 13, no. 3, p. e67, Sep. 2011.
[18] G. S. Reen, M. Mohandas, and S. Venkatesan, "Decentralized patient centric e-health record management system using blockchain and IPFS," in *Proc. IEEE Conf. Inf. Commun. Technol.*, Dec. 2019, pp. 1–7.
[19] A. A. Nancy, D. Ravindran, P. M. D. R. Vincent, K. Srinivasan, and D. G. Reina, "IoT-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning," *Electronics*, vol. 11, no. 15, p. 2292, Jul. 2022.
[20] J. Azmi, M. Arif, M. T. Nafis, M. A. Alam, S. Tanweer, and G. Wang, "A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data," *Med. Eng. Phys.*, vol. 105, Jul. 2022, Art. no. 103825.
[21] G. H. Lee and S.-Y. Shin, "Federated learning on clinical benchmark data: Performance assessment," *J. Med. Internet Res.*, vol. 22, no. 10, Oct. 2020, Art. no. e20891.
[22] P. Gangwani, A. Perez-Pons, S. Joshi, H. Upadhyay, and L. Lagos, "Integration of data science and IoT with blockchain for Industry 4.0," in *Blockchain and its Applications in Industry 4.0*. Cham, Switzerland: Springer, 2023, pp. 139–177.
[23] W. Zhao, I. M. Aldyaflah, P. Gangwani, S. Joshi, H. Upadhyay, and L. Lagos, "A blockchain-facilitated secure sensing data processing and logging system," *IEEE Access*, vol. 11, pp. 21712–21728, 2023.
[24] Q. Dou, "Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study," *NPJ Digit. Med.*, vol. 4, no. 1, p. 60, Mar. 2021.
[25] T. Benil and J. Jasper, "Cloud based security on outsourcing using blockchain in E-health systems," *Comput. Netw.*, vol. 178, Sep. 2020, Art. no. 107344.
[26] A. Razzaq, S. A. H. Mohsan, S. A. K. Ghayyur, N. Al-Kahtani, H. K. Alkahtani, and S. M. Mostafa, "Blockchain in healthcare: A decentralized platform for digital health passport of COVID-19 based on vaccination and immunity certificates," *Healthcare*, vol. 10, no. 12, p. 2453, Dec. 2022.
[27] D. R. Shah, D. A. Dhawan, S. N. Shah, P. R. Shah, and S. Francis, "Panacea: A novel architecture for electronic health records system using blockchain and machine learning," in *Proc. 2nd Int. Conf. Adv. Electr., Comput., Commun. Sustain. Technol. (ICAECT)*, Apr. 2022, pp. 1–7.
[28] L. Peng, N. Wang, N. Dvornek, X. Zhu, and X. Li, "FedNI: Federated graph learning with network inpainting for population-based disease prediction," *IEEE Trans. Med. Imag.*, vol. 42, no. 7, pp. 2032–2043, Jul. 2023.
[29] J. Jayabalan and N. Jeyanthi, "Scalable blockchain model using off-chain IPFS storage for healthcare data security and privacy," *J. Parallel Distrib. Comput.*, vol. 164, pp. 152–167, Jun. 2022.
[30] V. S. Parekh, S. Lai, V. Braverman, J. Leal, S. Rowe, J. J. Pillai, and M. A Jacobs, "Cross-domain federated learning in medical imaging," 2021, *arXiv:2112.10001*.
[31] S. Chenthara, K. Ahmed, H. Wang, and F. Whittaker, "Security and privacy-preserving challenges of e-Health solutions in cloud computing," *IEEE Access*, vol. 7, pp. 74361–74382, 2019.
[32] H. Qiu, M. Qiu, M. Liu, and G. Memmi, "Secure health data sharing for medical cyber-physical systems for the healthcare 4.0," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 9, pp. 2499–2505, Sep. 2020.
[33] B. Han, R. H. Jhaveri, H. Wang, D. Qiao, and J. Du, "Application of robust zero-watermarking scheme based on federated learning for securing the healthcare data," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 2, pp. 804–813, Feb. 2023.
[34] G. Al-Sumaidaee, R. Alkhudary, Z. Zilic, and A. Swidan, "Performance analysis of a private blockchain network built on hyperledger fabric for healthcare," *Inf. Process. Manage.*, vol. 60, no. 2, Mar. 2023, Art. no. 103160.
[35] M. Abouali, K. Sharma, O. Ajayi, and T. Saadawi, "Performance evaluation of secured blockchain-based patient health records sharing framework," in *Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS)*, Jun. 2022, pp. 1–7.
[36] A. Rehman, S. Abbas, M. A. Khan, T. M. Ghazal, K. M. Adnan, and A. Mosavi, "A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106019.
[37] X. Yang, T. Li, W. Xi, A. Chen, and C. Wang, "A blockchain-assisted verifiable outsourced attribute-based signcryption scheme for EHRs sharing in the cloud," *IEEE Access*, vol. 8, pp. 170713–170731, 2020.
[38] H. D. Hoang, D. T. T. Hien, T. C. Nhut, P. D. T. Quyen, P. T. Duy, and V.-H. Pham, "A blockchain-based secured and privacy-preserved personal healthcare record exchange system," in *Proc. IEEE Int. Conf. Mach. Learn. Appl. Netw. Technol. (ICMLANT)*, Dec. 2021, pp. 1–5.
[39] A. Kofod-Petersen, "How to do a structured literature review in computer science," *Version 0.1*, vol. 1, pp. 1–7, Oct. 2012.
[40] V. S. Marichamy and V. Natarajan, "Blockchain based securing medical records in big data analytics," *Data Knowl. Eng.*, vol. 144, Mar. 2023, Art. no. 102122.
[41] G. Zhang, X. Zhang, M. Bilal, W. Dou, X. Xu, and J. J. P. C. Rodrigues, "Identifying fraud in medical insurance based on blockchain and deep learning," *Future Gener. Comput. Syst.*, vol. 130, pp. 140–154, May 2022.
[42] T. Veeramakali, R. Siva, B. Sivakumar, P. C. S. Mahesh, and N. Krishnaraj, "An intelligent Internet of Things-based secure healthcare framework using blockchain technology with an optimal deep learning model," *J. Supercomput.*, vol. 77, no. 9, pp. 9576–9596, Sep. 2021.
[43] S. Singh, S. Rathore, O. Alfarraj, A. Tolba, and B. Yoon, "A framework for privacy-preservation of IoT healthcare data using federated learning and blockchain technology," *Future Gener. Comput. Syst.*, vol. 129, pp. 380–388, Apr. 2022.

[44] P. Kumar, R. Kumar, A. Kumar, A. A. Franklin, S. Garg, and S. Singh, "Blockchain and deep learning for secure communication in digital twin empowered industrial IoT network," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 5, pp. 2802–2813, Sep./Oct. 2022.

[45] B. Mallikarjuna, G. Shrivastava, and M. Sharma, "Blockchain technology: A DNN token-based approach in healthcare and COVID-19 to generate extracted data," *Expert Syst.*, vol. 39, no. 3, Mar. 2022, Art. no. e12778.

[46] G. N. Nguyen, N. H. L. Viet, M. Elhoseny, K. Shankar, B. B. Gupta, and A. A. A. El-Latif, "Secure blockchain enabled Cyber–physical systems in healthcare using deep belief network with ResNet model," *J. Parallel Distrib. Comput.*, vol. 153, pp. 150–160, Jul. 2021.

[47] F. Ni, J. Zhang, and M. N. Noori, "Deep learning for data anomaly detection and data compression of a long-span suspension bridge," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 35, no. 7, pp. 685–700, Jul. 2020.

[48] V. Hassija, V. Gupta, S. Garg, and V. Chamola, "Traffic jam probability estimation based on blockchain and deep neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3919–3928, Jul. 2021.

[49] A. Ali, M. F. Pasha, J. Ali, O. H. Fang, M. Masud, A. D. Jurcut, and M. A. Alzain, "Deep learning based homomorphic secure search-able encryption for keyword search in blockchain healthcare system: A novel approach to cryptography," *Sensors*, vol. 22, no. 2, p. 528, Jan. 2022.

[50] M. Hassan, J. Chen, C. Zhu, and U. Zukaib, "Adoption of blockchain-based artificial intelligence in healthcare," in *Proc. 5th Int. Conf. Artif. Intell. Big Data (ICAIBD)*, May 2022, pp. 140–144.

[51] M. M. Madine, A. A. Battah, I. Yaqoob, K. Salah, R. Jayaraman, Y. Al-Hammadi, S. Pesic, and S. Ellahham, "Blockchain for giving patients control over their medical records," *IEEE Access*, vol. 8, pp. 193102–193115, 2020.

[52] B. Shen, J. Guo, and Y. Yang, "MedChain: Efficient healthcare data sharing via blockchain," *Appl. Sci.*, vol. 9, no. 6, p. 1207, Mar. 2019.

[53] J. Wiens, S. Saria, M. Sendak, M. Ghassemi, V. X. Liu, F. Doshi-Velez, K. Jung, K. Heller, D. Kale, M. Saeed, P. N. Ossorio, S. Thadaney-Israni, and A. Goldenberg, "Do no harm: A roadmap for responsible machine learning for health care," *Nature Med.*, vol. 25, no. 9, pp. 1337–1340, Sep. 2019.

[54] A. Akhtar, "Prediction of COVID-19 using ensemble based machine learning approach," *Int. J. Comput. Innov. Sci.*, vol. 1, no. 4, pp. 34–41, 2022.

[55] P. Ghosh, S. Azam, M. Jonkman, A. Karim, F. M. J. M. Shamrat, E. Ignatious, S. Shultana, A. R. Beeravolu, and F. De Boer, "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques," *IEEE Access*, vol. 9, pp. 19304–19326, 2021.

[56] A. Ghasemieh, A. Lloyed, P. Bahrami, P. Vajar, and R. Kashef, "A novel machine learning model with stacking ensemble learner for predicting emergency readmission of heart-disease patients," *Decis. Anal. J.*, vol. 7, Jun. 2023, Art. no. 100242.

[57] I. Sardar, M. A. Akbar, V. Leiva, A. Alsanad, and P. Mishra, "Machine learning and automatic ARIMA/prophet models-based forecasting of COVID-19: Methodology, evaluation, and case study in SAARC countries," *Stochastic Environ. Res. Risk Assessment*, vol. 37, no. 1, pp. 345–359, Jan. 2023.

[58] Z. Ibrahim, P. Tulay, and J. Abdullahi, "Multi-region machine learning-based novel ensemble approaches for predicting COVID-19 pandemic in Africa," *Environ. Sci. Pollut. Res.*, vol. 30, no. 2, pp. 3621–3643, Jan. 2023.

[59] M. A. Khan, "An IoT framework for heart disease prediction based on MDCNN classifier," *IEEE Access*, vol. 8, pp. 34717–34727, 2020.

[60] F. Ali, S. El-Sappagh, S. M. R. Islam, D. Kwak, A. Ali, M. Imran, and K.-S. Kwak, "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion," *Inf. Fusion*, vol. 63, pp. 208–222, Nov. 2020.

[61] D. Zhang, Y. Chen, Y. Chen, S. Ye, W. Cai, J. Jiang, Y. Xu, G. Zheng, and M. Chen, "Heart disease prediction based on the embedded feature selection method and deep neural network," *J. Healthcare Eng.*, vol. 2021, pp. 1–9, Sep. 2021.

[62] A. Gupta, R. Kumar, H. S. Arora, and B. Raman, "C-CADZ: Computational intelligence system for coronary artery disease detection using Z-Alizadeh sani dataset," *Int. J. Speech Technol.*, vol. 52, no. 3, pp. 2436–2464, Feb. 2022.

[63] R. Krishnamoorthi, S. Joshi, H. Z. Almarzouki, P. K. Shukla, A. Rizwan, C. Kalpana, and B. Tiwari, "A novel diabetes healthcare disease prediction framework using machine learning techniques," *J. Healthcare Eng.*, vol. 2022, pp. 1–10, Jan. 2022.

[64] M. De Bois, M. A. El Yacoubi, and M. Ammi, "Enhancing the interpretability of deep models in healthcare through attention: Application to glucose forecasting for diabetic people," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 35, no. 12, Sep. 2021, Art. no. 2160006.

[65] H. Li, C. Li, J. Wang, A. Yang, Z. Ma, Z. Zhang, and D. Hua, "Review on security of federated learning and its application in healthcare," *Future Gener. Comput. Syst.*, vol. 144, pp. 271–290, Jul. 2023.

[66] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, "Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records," *J. Biomed. Informat.*, vol. 99, Nov. 2019, Art. no. 103291.

[67] H. Gao, N. He, and T. Gao, "SVeriFL: Successive verifiable federated learning with privacy-preserving," *Inf. Sci.*, vol. 622, pp. 98–114, Apr. 2023.

[68] D. J. Beutel, T. Topal, A. Mathur, X. Qiu, J. Fernandez-Marques, Y. Gao, L. Sani, K. H. Li, T. Parcollet, P. P. B. de Gusmão, and N. D. Lane, "Flower: A friendly federated learning research framework," 2020, *arXiv:2007.14390*.

[69] T.-L. Yang, H.-W. Tsai, W.-C. Huang, J.-C. Lin, J.-B. Liao, N.-H. Chow, and P.-C. Chung, "Pathologic liver tumor detection using feature aligned multi-scale convolutional network," *Artif. Intell. Med.*, vol. 125, Mar. 2022, Art. no. 102244.

[70] N. Alalwan, A. Abozeid, A. A. ElHabshy, and A. Alzahrani, "Efficient 3D deep learning model for medical image semantic segmentation," *Alexandria Eng. J.*, vol. 60, no. 1, pp. 1231–1239, Feb. 2021.

[71] N. Cherukuri, N. R. Bethapudi, V. S. K. Thotakura, P. Chitturi, C. Z. Basha, and R. M. Mummidi, "Deep learning for lung cancer prediction using NSCLS patients CT information," in *Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS)*, Mar. 2021, pp. 325–330.

[72] D. Gao, C. Ju, X. Wei, Y. Liu, T. Chen, and Q. Yang, "HHHFL: Hierarchical heterogeneous horizontal federated learning for electroencephalography," 2019, *arXiv:1909.05784*.

[73] D. Karimi, G. Nir, L. Fazli, P. C. Black, L. Goldenberg, and S. E. Salcudean, "Deep learning-based Gleason grading of prostate cancer from histopathology images—Role of multiscale decision aggregation and data augmentation," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 5, pp. 1413–1426, May 2020.

[74] Y. Chang, H. Park, H.-J. Yang, S. Lee, K.-Y. Lee, T. S. Kim, J. Jung, and J.-M. Shin, "Cancer drug response profile scan (CDRscan): A deep learning model that predicts drug effectiveness from cancer genomic signature," *Sci. Rep.*, vol. 8, no. 1, p. 8857, Jun. 2018.

[75] T. Borger, P. Mosteiro, H. Kaya, E. Rijcken, A. A. Salah, F. Scheepers, and M. Spruit, "Federated learning for violence incident prediction in a simulated cross-institutional psychiatric setting," *Expert Syst. Appl.*, vol. 199, Aug. 2022, Art. no. 116720.

[76] X. Xu, H. Peng, L. Sun, M. Z. A. Bhuiyan, L. Liu, and L. He, "FedMood: Federated learning on mobile health data for mood detection," 2021, *arXiv:2102.09342*.

[77] J. Lo, T. T. Yu, D. Ma, P. Zang, J. P. Owen, Q. Zhang, R. K. Wang, M. F. Beg, A. Y. Lee, Y. Jia, and M. V. Sarunic, "Federated learning for microvasculature segmentation and diabetic retinopathy classification of OCT data," *Ophthalmol. Sci.*, vol. 1, no. 4, Dec. 2021, Art. no. 100069.

[78] J. Cheng, P. Luo, N. Xiong, and J. Wu, "AAFL: Asynchronous-adaptive federated learning in edge-based wireless communication systems for countering communicable infectious diseasess," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 11, pp. 3172–3190, Nov. 2022.

[79] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R. R. Colen, and S. Bakas, "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," *Sci. Rep.*, vol. 10, no. 1, Jul. 2020, Art. no. 12598.

[80] D. Yang, Z. Xu, W. Li, A. Myronenko, H. R. Roth, S. Harmon, S. Xu, B. Turkbey, E. Turkbey, X. Wang, W. Zhu, G. Carrafiello, F. Patella, M. Cariati, H. Obinata, H. Mori, K. Tamura, P. An, B. J. Wood, and D. Xu, "Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan," *Med. Image Anal.*, vol. 70, May 2021, Art. no. 101992.

[81] I. Feki, S. Ammar, Y. Kessentini, and K. Muhammad, "Federated learning for COVID-19 screening from chest X-ray images," *Appl. Soft Comput.*, vol. 106, Jul. 2021, Art. no. 107330.
[82] C. I. Bercea, B. Wiestler, D. Rueckert, and S. Albarqouni, "Federated disentangled representation learning for unsupervised brain anomaly detection," *Nature Mach. Intell.*, vol. 4, no. 8, pp. 685–695, Aug. 2022.
[83] U. Ahmed, J. C. Lin, and G. Srivastava, "Hyper-graph attention based federated learning methods for use in mental health detection," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 2, pp. 768–777, Feb. 2023.
[84] A. Vaid, "Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: Machine learning approach," *JMIR Med. Informat.*, vol. 9, no. 1, Jan. 2021, Art. no. e24207.
[85] I. Dayan, "Federated learning for predicting clinical outcomes in patients with COVID-19," *Nature Med.*, vol. 27, no. 10, pp. 1735–1743, 2021.
[86] A. Sadilek, L. Liu, D. Nguyen, M. Kamruzzaman, S. Serghiou, B. Rader, A. Ingerman, S. Mellem, P. Kairouz, and E. O. Nsoesie, "Privacy-first health research with federated learning," *NPJ Digit. Med.*, vol. 4, no. 1, p. 132, 2021.
[87] Y.-L. Huang, H.-C. Yang, and C.-C. Lee, "Federated learning via conditional mutual learning for Alzheimer's disease classification on T1w MRI," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 2427–2432.
[88] H. Lee, Y. J. Chai, H. Joo, K. Lee, J. Y. Hwang, S.-M. Kim, K. Kim, I.-C. Nam, J. Y. Choi, H. W. Yu, M.-C. Lee, H. Masuoka, A. Miyauchi, K. E. Lee, S. Kim, and H.-J. Kong, "Federated learning for thyroid ultrasound image analysis to protect personal information: Validation study in a real health care environment," *JMIR Med. Informat.*, vol. 9, no. 5, May 2021, Art. no. e25869.
[89] M. Y. Lu, R. J. Chen, D. Kong, J. Lipkova, R. Singh, D. F. K. Williamson, T. Y. Chen, and F. Mahmood, "Federated learning for computational pathology on gigapixel whole slide images," *Med. Image Anal.*, vol. 76, Feb. 2022, Art. no. 102298.
[90] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," Softw. Eng. Group School Comput. Sci. Math., Dept. Comput. Sci. Univ. Durham, Keele Univ., Keele, U.K., Durham, U.K., Tech. Rep., 2007.
[91] D. Moher, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Ann. Internal Med.*, vol. 151, no. 4, p. 264, Aug. 2009.
[92] G. Hegde, S. Bekal, S. S. Prasad, P. D. Shenoy, and K. R. Venugopal, "Analysis of secure EHR storage methods on blockchain and integrating ML to predict chronic kidney disease," in *Proc. IEEE 7th Int. Conf. Recent Adv. Innov. Eng. (ICRAIE)*, vol. 7, Dec. 2022, pp. 250–255.
[93] V. Patel, "A framework for secure and decentralized sharing of medical imaging data via blockchain consensus," *Health Informat. J.*, vol. 25, no. 4, pp. 1398–1411, Dec. 2019.
[94] Y.-S. Bae, Y. Park, T. Kim, T. Ko, M.-S. Kim, E. Lee, H.-C. Kim, and H.-J. Yoon, "Development and pilot-test of blockchain-based MyHealthData platform," *Appl. Sci.*, vol. 11, no. 17, p. 8209, Sep. 2021.
[95] R. Guo, H. Shi, Q. Zhao, and D. Zheng, "Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems," *IEEE Access*, vol. 6, pp. 11676–11686, 2018.
[96] F. Tang, S. Ma, Y. Xiang, and C. Lin, "An efficient authentication scheme for blockchain-based electronic health records," *IEEE Access*, vol. 7, pp. 41678–41689, 2019.
[97] O. Gutiérrez, G. Romero, L. Pérez, A. Salazar, M. Charris, and P. Wightman, "HealthyBlock: Blockchain-based IT architecture for electronic medical records resilient to connectivity failures," *Int. J. Environ. Res. Public Health*, vol. 17, no. 19, p. 7132, Sep. 2020.
[98] A. Martínez, C. Molina, and D. Subauste, "Electronic medical records management in health organizations using a technology architecture based on blockchain," in *Proc. IEEE ANDESCON*, Oct. 2020, pp. 1–6.
[99] A. Donawa, I. Orukari, and C. E. Baker, "Scaling blockchains to support electronic health records for hospital systems," in *Proc. IEEE 10th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2019, pp. 550–556.
[100] A. Shahnaz, U. Qamar, and A. Khalid, "Using blockchain for electronic health records," *IEEE Access*, vol. 7, pp. 147782–147795, 2019.
[101] Y. Wang and M. He, "CPDS: A cross-blockchain based privacy-preserving data sharing for electronic health records," in *Proc. IEEE 6th Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Apr. 2021, pp. 90–99.
[102] S. Jiang, H. Wu, and L. Wang, "Patients-controlled secure and privacy-preserving EHRs sharing scheme based on consortium blockchain," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
[103] N. Sammeta and L. Parthiban, "Hyperledger blockchain enabled secure medical record management with deep learning-based diagnosis model," *Complex Intell. Syst.*, vol. 8, no. 1, pp. 625–640, Feb. 2022.
[104] X. Zhang and S. Poslad, "Blockchain support for flexible queries with granular access control to electronic medical records (EMR)," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
[105] H. Guo, W. Li, M. Nejad, and C.-C. Shen, "Access control for electronic health records with hybrid blockchain-edge architecture," in *Proc. IEEE Int. Conf. Blockchain (Blockchain)*, Jul. 2019, pp. 44–51.
[106] A. R. Rajput, Q. Li, M. T. Ahvanooey, and I. Masood, "EACMS: Emergency access control management system for personal health record based on blockchain," *IEEE Access*, vol. 7, pp. 84304–84317, 2019.
[107] R. K. Marangappanavar and M. Kiran, "Inter-planetary file system enabled blockchain solution for securing healthcare records," in *Proc. 3rd ISEA Conf. Secur. Privacy (ISEA-ISAP)*, Feb. 2020, pp. 171–178.
[108] X. Wu, Y. Han, M. Zhang, and S. Zhu, "Secure personal health records sharing based on blockchain and IPFS," in *Proc. Chin. Conf. Trusted Comput. Inf. Secur.* Shanghai, China, Cham, Switzerland: Springer, Oct. 2019, pp. 340–354.
[109] J. Xu, K. Xue, S. Li, H. Tian, J. Hong, P. Hong, and N. Yu, "Healthchain: A blockchain-based privacy preserving scheme for large-scale health data," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8770–8781, Oct. 2019.
[110] R. Jabbar, N. Fetais, M. Krichen, and K. Barkaoui, "Blockchain technology for healthcare: Enhancing shared electronic health record interoperability and integrity," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 310–317.
[111] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Blockchain for secure EHRs sharing of mobile cloud based E-Health systems," *IEEE Access*, vol. 7, pp. 66792–66806, 2019.
[112] A. Dubovitskaya, F. Baig, Z. Xu, R. Shukla, P. S. Zambani, A. Swaminathan, M. M. Jahangir, K. Chowdhry, R. Lachhani, N. Idnani, M. Schumacher, K. Aberer, S. D. Stoller, S. Ryu, and F. Wang, "ACTION-EHR: Patient-centric blockchain-based electronic health record data management for cancer care," *J. Med. Internet Res.*, vol. 22, no. 8, Aug. 2020, Art. no. e13598.
[113] P. Chinnasamy and P. Deepalakshmi, "HCAC-EHR: Hybrid cryptographic access control for secure EHR retrieval in healthcare cloud," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 2, pp. 1001–1019, Feb. 2022.
[114] A. Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, and F. Wang, "Secure and trustable electronic medical records sharing using blockchain," in *Proc. AMIA Annu. Symp.*, 2017, p. 650.
[115] B. Vardhini, S. N. Dass, S. R, and R. Chinnaiyan, "A blockchain based electronic medical health records framework using smart contracts," in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2021, pp. 1–4.
[116] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using blockchain for medical data access and permission management," in *Proc. 2nd Int. Conf. Open Big Data (OBD)*, Aug. 2016, pp. 25–30.
[117] G. Carter, H. Shahriar, and S. Sneha, "Blockchain-based interoperable electronic health record sharing framework," in *Proc. IEEE 43rd Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, vol. 2, Jul. 2019, pp. 452–457.
[118] H. H. Kung, Y.-F. Cheng, H.-A. Lee, and C.-Y. Hsu, "Personal health record in FHIR format based on blockchain architecture," in *Proc. Int. Conf. Frontier Comput.* Cham, Switzerland: Springer, 2020, pp. 1776–1788.
[119] A. Tang, J. Han, and P. Chen, "A comparative analysis of architecture frameworks," in *Proc. 11th Asia–Pacific Softw. Eng. Conf.*, Nov. 2004, pp. 640–647.
[120] M. Biehl, *API Architecture*, vol. 2. Brisbane, QLD, Australia: API-Univ. Press, 2015.
[121] M. Hamdaqa, L. A. P. Met, and I. Qasse, "IContractML 2.0: A domain-specific language for modeling and deploying smart contracts onto multiple blockchain platforms," *Inf. Softw. Technol.*, vol. 144, Apr. 2022, Art. no. 106762.
[122] R. A. de Páez, "Un acercamiento a la reutilización en ingeniería de software," *Revista Universidad EAFIT*, vol. 35, no. 114, pp. 51–63, 1999.
[123] A. A. Mamun, S. Azam, and C. Gritti, "Blockchain-based electronic health records management: A comprehensive review and future research direction," *IEEE Access*, vol. 10, pp. 5768–5789, 2022.

[124] P. Zhang, J. White, D. C. Schmidt, G. Lenz, and S. T. Rosenbloom, "FHIRChain: Applying blockchain to securely and scalably share clinical data," *Comput. Structural Biotechnol. J.*, vol. 16, pp. 267–278, Jan. 2018.

[125] H. Jin, C. Xu, Y. Luo, and P. Li, "Blockchain-based secure and privacy-preserving clinical data sharing and integration," in *Proc. Int. Conf. Algorithms Archit. Parallel Process.* New York, NY, USA, Oct. 2020, pp. 93–109.

[126] B Arunkumar and G Kousalya, "Blockchain-based decentralized and secure lightweight e-health system for electronic health records," in *Intelligent Systems, Technologies and Applications*. Cham, Switzerland: Springer, 2020, pp. 273–289.

[127] A. K. Bashir, N. Victor, S. Bhattacharya, T. Huynh-The, R. Chengoden, G. Yenduri, P. K. R. Maddikunta, Q.-V. Pham, T. R. Gadekallu, and M. Liyanage, "Federated learning for the healthcare metaverse: Concepts, applications, challenges, and future directions," *IEEE Internet Things J.*, early access, Aug. 14, 2023, doi: 10.1109/JIOT.2023.3304790.

[128] G. Nagasubramanian, R. K. Sakthivel, R. Patan, A. H. Gandomi, M. Sankayya, and B. Balusamy, "Securing e-health records using keyless signature infrastructure blockchain technology in the cloud," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 639–647, Feb. 2020.

[129] D. Tith, J.-S. Lee, H. Suzuki, W. M. A. B. Wijesundara, N. Taira, T. Obi, and N. Ohyama, "Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability," *Healthcare Informat. Res.*, vol. 26, no. 1, p. 3, 2020.

[130] A. Haddad, M. H. Habaebi, M. R. Islam, N. F. Hasbullah, and S. A. Zabidi, "Systematic review on AI-blockchain based E-Healthcare records management systems," *IEEE Access*, vol. 10, pp. 94583–94615, 2022.

[131] P. Meier, J. H. Beinke, C. Fitte, J. Schulte to Brinke, and F. Teuteberg, "Generating design knowledge for blockchain-based access control to personal health records," *Inf. Syst. e-Bus. Manage.*, vol. 19, no. 1, pp. 13–41, Mar. 2021.

[132] A. A. Mazlan, S. M. Daud, S. M. Sam, H. Abas, S. Z. A. Rasid, and M. F. Yusof, "Scalability challenges in healthcare blockchain system—A systematic review," *IEEE Access*, vol. 8, pp. 23663–23673, 2020.

[133] L. Ismail, H. Materwala, and S. Zeadally, "Lightweight blockchain for healthcare," *IEEE Access*, vol. 7, pp. 149935–149951, 2019.

[134] Z. K. Taha, C. T. Yaw, S. P. Koh, S. K. Tiong, K. Kadirgama, F. Benedict, J. D. Tan, and Y. A. Balasubramaniam, "A survey of federated learning from data perspective in the healthcare domain: Challenges, methods, and future directions," *IEEE Access*, vol. 11, pp. 45711–45735, 2023.

[135] A. A. Omar, M. Z. A. Bhuiyan, A. Basu, S. Kiyomoto, and M. S. Rahman, "Privacy-friendly platform for healthcare data in cloud based on blockchain environment," *Future Gener. Comput. Syst.*, vol. 95, pp. 511–521, Jun. 2019.

[136] C. Zhao, S. Zhao, M. Zhao, Z. Chen, C.-Z. Gao, H. Li, and Y.-A. Tan, "Secure multi-party computation: Theory, practice and applications," *Inf. Sci.*, vol. 476, pp. 357–372, Feb. 2019.

[137] E. Badidi, K. Moumane, and F. E. Ghazi, "Opportunities, applications, and challenges of edge-AI enabled video analytics in smart cities: A systematic review," *IEEE Access*, vol. 11, pp. 80543–80572, 2023.

[138] Y. Chang, C. Fang, and W. Sun, "A blockchain-based federated learning method for smart healthcare," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–12, Nov. 2021.

[139] O. Vandenberg, D. Martiny, O. Rochas, A. van Belkum, and Z. Kozlakidis, "Considerations for diagnostic COVID-19 tests," *Nature Rev. Microbiol.*, vol. 19, no. 3, pp. 171–183, Mar. 2021.

[140] H. B. Mahajan, A. S. Rashid, A. A. Junnarkar, N. Uke, S. D. Deshpande, P. R. Futane, A. Alkhayyat, and B. Alhayani, "Integration of healthcare 4.0 and blockchain into secure cloud-based electronic health records systems," *Appl. Nanoscience*, vol. 13, no. 3, pp. 2329–2342, 2023.

[141] M. Hassan, C. Jincai, A. Iftekhar, and X. Cui, "Future of the Internet of Things emerging with blockchain and smart contracts," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 6, pp. 1–5, 2020.

[142] A. Kumari, R. Gupta, S. Tanwar, and N. Kumar, "Blockchain and AI amalgamation for energy cloud management: Challenges, solutions, and future directions," *J. Parallel Distrib. Comput.*, vol. 143, pp. 148–166, Sep. 2020.

[143] X. Liu, Z. Wang, C. Jin, F. Li, and G. Li, "A blockchain-based medical data sharing and protection scheme," *IEEE Access*, vol. 7, pp. 118943–118953, 2019.

[144] M. S. Rahman, I. Khalil, P. C. M. Arachchige, A. Bouras, and X. Yi, "A novel architecture for tamper proof electronic health record management system using blockchain wrapper," in *Proc. ACM Int. Symp. Blockchain Secure Crit. Infrastruct.*, Jul. 2019, pp. 97–105.

[145] Y. Teng, Y. Cao, M. Liu, F. R. Yu, and V. C. M. Leung, "Efficient blockchain-enabled large scale parked vehicular computing with green energy supply," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9423–9436, Sep. 2021.

[146] H. F. Anjum, S. Z. A. Rasid, H. Khalid, M. M. Alam, S. M. Daud, H. Abas, S. M. Sam, and M. F. Yusof, "Mapping research trends of blockchain technology in healthcare," *IEEE Access*, vol. 8, pp. 174244–174254, 2020.

[147] S. Khatri, F. A. Alzahrani, M. T. J. Ansari, A. Agrawal, R. Kumar, and R. A. Khan, "A systematic analysis on blockchain integration with healthcare domain: Scope and challenges," *IEEE Access*, vol. 9, pp. 84666–84687, 2021.

[148] T. Nusairat, M. M. Saudi, and A. B. Ahmad, "A recent assessment for the ransomware attacks against the Internet of Medical Things (IoMT): A review," in *Proc. IEEE 13th Int. Conf. Control Syst., Comput. Eng. (ICCSCE)*, Aug. 2023, pp. 238–242.

[149] J. Alenizi and I. Alrashdi, "SFMR-SH: Secure framework for mitigating ransomware attacks in smart healthcare using blockchain technology," *Sustain. Mach. Intell. J.*, vol. 2, pp. 1–19, Mar. 2023.

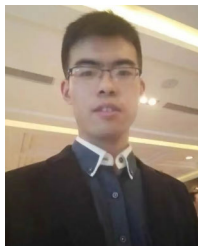
**UMER ZUKAIB** is currently pursuing the Ph.D. degree in cyberspace security with the School of Cyber Science and Engineering, Wuhan University, China. His current research interests include threat detection, monitoring, mitigation, blockchain, and machine learning.

**XIAOHUI CUI** received the Ph.D. degree in computer science and engineering from the University of Louisville, Louisville, KY, USA, in 2004. He is currently a Professor with the School of Cyber Science and Engineering, Wuhan University. His research interests include artificial intelligence, blockchain technology, and high-performance computing.

**MIR HASSAN** (Member, IEEE) is currently pursuing the Ph.D. degree in artificial intelligence with the Department of Information Engineering and Computer Science, University of Trento, Italy. He is also working on the EIC-funded SUST(AI)N project for the unparalleled development of precision-sensing AI processing shared across devices (i.e., distributed intelligence) and the flexibility and implementation efficiency of reconfigurable hardware. He was a Specialist in the BIO-MAC Project with Vilnius Tech University, Vilnius, Lithuania. His current research interests include federated learning, blockchain technology, distributed intelligence, embedded systems, and the Internet of Things.

**SHEETAL HARRIS** is currently pursuing the Ph.D. degree in cyberspace security with the School of Cyber Science and Engineering, Wuhan University, China. Her current research interests include fake news detection, NLP, and machine learning.

**HASSAN JALIL HADI** received the M.S. degree in information security from Riphah International University, Islamabad, Pakistan. He is currently pursuing the Ph.D. degree in cyberspace security with the School of Cyber Science and Engineering, Wuhan University, China. He is also a Cyber-Security Expert performing threat detection, monitoring, mitigation, and audit on industry-proven standards for CS challenges with the Cyber Reconnaissance and Combat (CRC) Laboratory, National Center for Cyber Security, Islamabad. His research interests include the IoT security, digital forensics, edge computing, and botnet. Beyond his academic achievements and research endeavors, he holds credentials from ISACA and EC-Council, including the Certified Information Security Manager (CISM) for strategic security management. He has certified in CEH, CHFI, and ISO 27001, showcasing expertise in ethical hacking, forensics, and information security management.



**CHENGLIANG ZHENG** is currently pursuing the Ph.D. degree in cyberspace security with the School of Cyber Science and Engineering, Wuhan University, China. His current research interests include blockchain and machine learning.