



Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/combiomed

Self-sovereign identity empowered non-fungible patient tokenization for health information exchange using blockchain technology

Yan Zhuang^{a,b}, Chi-Ren Shyu^d, Shenda Hong^{a,b}, Pengfei Li^{a,b,c}, Luxia Zhang^{a,b,c,*}

^a National Institute of Health Data Science, Peking University, Beijing, China

^b Institute of Medical Technology, Health Science Center of Peking University, Beijing, China

^c Advanced Institute of Information Technology, Peking University, Hangzhou, China

^d Institute of Data Science and Informatics, Columbia, MO, USA

ARTICLE INFO

Keywords:

Blockchain
Self-sovereign identity
Non-fungible token
Health information exchange
Patient tokenization
Smart contract

ABSTRACT

Background: Patient tokenization is a novel approach that allows anonymous patient-level linkage across healthcare facilities, minimizing the risk of breaching protected health information in health information exchange (HIE). Most patient tokenization is the centralized approach that is unable to address data security concerns fundamentally. Non-Fungible Tokens (NFT), which are non-transferable cryptographic assets on the blockchain, have the potential to provide secure, decentralized, and trustworthy patient tokenization. Self-Sovereign Identity (SSI) is a user-centric approach to verify the ownership of NFTs in a decentralized manner. **Methods:** We have developed a blockchain architecture that contains four modules: (1) Creation module for NFTs creation, (2) Linkage module to link the local patients' accounts to their NFTs, (3) Authentication module that allows patients to permit healthcare providers to access their token, and (4) Exchange module, which involves the HIE process and the validation of the legitimacy of the token through SSI.

Results: A case study has been conducted on the proposed architecture. Over 3 million transactions have been completed successfully with a blockchain validation and written time of 1.17 s on average. A stability test has also been conducted with a higher throughput of 200 transactions per second running for an hour with an average transaction processing time of 1.42 s.

Conclusions: This study proposed a blockchain architecture that achieves SSI-enabled NFT-based patient tokenization. Our architecture design, implementation, and case studies have demonstrated the feasibility and potential of NFT with SSI to establish a secure, transparent, and patient-centric identity management and HIE.

1. Introduction

Patient privacy and confidentiality are part of the healthcare process and vital to the establishment of trust between patients and healthcare providers. Under the regulations specified by the Health Insurance Portability and Accountability Act (HIPAA), healthcare organizations must implement precautions to protect patients' protected health information (PHI), which include all individually identifiable information, from unauthorized data access or breaches [1,2]. Violations can result in hefty penalties and even jail time in severe cases [3]. However, the United States Department of Health and Human Services received reports of breaches from over 550 organizations in 2021, with over 40 million people affected, a roughly 20-fold increase over the previous year [4]. As data access during healthcare continues to grow in terms of

the volume and the variety of data sources thereby increasing the exposure of patient records, there is a pressing need to anonymously bridge data from disparate sources at the patient level [5], providing healthcare professionals and researchers with a comprehensive picture of the patient's medical history while minimizing the risk of PHI breaches [6].

The master patient index, which incorporates demographic information contained in the PHI, is used to identify patients across health systems during Health Information Exchange (HIE) [7], the process of sharing Electronic Health Records (EHR) among healthcare professionals. As third-party organizations execute the majority of today's HIE, this raises concerns about data breaches, undermines trust, and limits patients' capacity to decide what, when, and with whom to share their data [8,9]. Patient tokenization is a new approach of replacing an

* Corresponding author. National Institute of Health Data Science Peking University Health Science Center No.38 Xueyuan Rd, Haidian District, Beijing, China.
E-mail address: zhanglx@bjmu.edu.cn (L. Zhang).

<https://doi.org/10.1016/j.combiomed.2023.106778>

Received 20 November 2022; Received in revised form 30 January 2023; Accepted 9 March 2023

Available online 15 March 2023

0010-4825/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

individual's PHI with a unique identification value linking data from multiple healthcare organizations, even if the information does not match due to changes in personal information (e.g., name, address) or human errors from the input [10]. Patients can decide how their data is shared and store the access history via tokenization. However, in the current market, patient tokenization is also produced and managed by third-party organizations, thus this approach has not adequately addressed data breaches and security concerns [11]. Decentralized technology, such as blockchain, is perceived as a potential that may maximize the benefits of patient tokenization.

Blockchain is a decentralized, open-source, distributed ledger system formed by continuously generated blocks that contain all transactions that record the data requests and exchanges [12]. Unique features of blockchain, such as transparency (all data in the blockchain is publicly accessible and auditable), immutability (all data stored inside the blockchain is technically unchangeable), and a mechanism for consensus (transactions need to be verified by most users with respect to the legitimacy of the data provenance before being recorded in the blockchain), have contributed to the success of blockchain [13]. Many researchers have examined the potential uses of blockchain in healthcare tasks in terms of a range of aspects [12], including EHRs [14], clinical trials [15,16], and pharmaceutical supply chains [17]; these scholars have empirically proved the feasibility, stability, and security of the technology [18]. Cryptocurrencies were the first blockchain application, in the form of a digitalized currency that experienced extremely fast acceptance by the global market. Cryptocurrency tokens represent interchangeable digital assets and are created using smart contracts, which are self-executing programable protocols on the blockchain [19]. Different from traditional tokens or cryptocurrencies, NFTs are one-of-a-kind digital assets that cannot be duplicated or divided. Due to their inherent value, NFT can only be exchanged, but their value cannot be directly compared to other tokens [20]. The nature of NFTs is that, rather than carrying content information, they contain metadata that points to digital or physical artifacts and is stored in smart contracts, making the record tamper-proof, transparent, and secure. As a result, most NFTs are designed as ownership for tradeable digital assets, and the NFT market has ballooned to \$41 billion in 2021 [21]. PHI constitutes an asset to patients in the healthcare setting, thus NFT can be an anonymous patient token that connects them across healthcare facilities without exposing their PHI [22]. Furthermore, NFTs can achieve patient-centric data management, allowing patients to access, authenticate, and share any of their records with anyone at any time. A recent study published in Science offers a diverse range of viewpoints on the potential transformation of HIE through the use of NFTs to protect PHI [23]. To maintain the pseudonymity of NFT and blockchain, a mechanism to verify ownership of the NFT without revealing the patient's actual identity is required [24].

Another work published in Cell shows the importance of using Self-Sovereign Identity (SSI), a user-centric approach to giving individuals full control over and consent to having their identities managed in a decentralized manner [25,26]. In contrast to mainstream identity management models such as the centralized and federated model (a federation of centralized models) which requires users to be registered with a third-party login service using credentials stored and verified by a centralized entity, SSI credentials are created using cryptography, i.e., a unique pair of public and private keys that keeps users pseudo-anonymous and that can be verified anywhere at any time through blockchain [27]. For example, an NFT connected to the patient's PHI and signed by both the issuer's private key as well as the owner's public key is given to a patient by a healthcare facility. When the patient visits another healthcare facility and desires HIE to access their medical records, they may present and authorize the NFT to the second healthcare provider. The verifier can validate the NFT's ownership by matching the patient's signature to the patient's public key stored in the NFT, as well as the data source's authenticity by comparing the issuer's public key to the issuer's signature in the NFT [20]. Without

the requirement for a third-party intermediary, the HIE can proceed directly between these two healthcare organizations. Blockchain-based SSI adopts the "Zero-Knowledge Proof" concept [28], which means the verifier will have "zero-knowledge" of the NFT but will be convinced of its authenticity (the "proof") without having any readable information inside the NFT to support the proof, to establish trust in peer-to-peer channels between the issuer, owner, and verifier, also known as the "trust triangle" [29].

The remaining sections of this paper are organized as follows: Section 2 surveys the existing blockchain-based research approaches for decentralized identity and compares our research with related work; Section 3 details the design and implementation of the four modules of blockchain-based architecture for NFT *creation, linkage* from multiple healthcare facilities, patient *authentication, exchange* to a healthcare provider, and SSI-based validation by the blockchain; Section 4 describes a case study that mimics the NFT authentication and exchange processes and analyzes the results to assess the system's feasibility and performance; and Section 5 presents remarks of our key contributions, limitations, and future research directions on extending the blockchain architecture for patient tokenization.

2. Related work

While the majority of blockchain research in healthcare focuses on patient-centered interoperable HIE, several eminent scholars have proposed blockchain approaches for decentralized identity in general rather than healthcare [30]. In this section, we outline related prominent works [31,32] and compare them with our proposed blockchain platform for patient tokenization (see Table 1).

ShoCard [33] is one of the earliest enterprise solutions for users to keep and protect their identities using blockchain. Users' trusted credentials are stored locally while linking to the cryptographic hash in the blockchain transaction to achieve SSI. ShoCard provides multi-factor authentication where users can log in to the blockchain service without a password. However, it requires the centralized ShoCard server's storage of encrypted sensitive data, such as biometric data while enrolling in the system. Although the central server provides security mechanisms for data access and sharing, these settings nonetheless raise concerns about the leakage of users' sensitive data and the potential risk that users will not be able to revoke their data if the server terminates.

uPort [34] is the first identity system to support SSI, letting the user have full control over their identity built on the Ethereum blockchain by ConsenSys, one of the leading blockchain startup companies. Users can generate an asymmetric key pair through their mobile app and transmit a transaction to Ethereum to keep a reference to their account. uPort supports key recovery by nominating trustees, voting through the smart contract, and generating new pair of keys but keeping the original ID. There is no centralized server to authenticate users; this eliminates the risk of data breaches from third parties. However, since the registry smart contract stores all of the registering attributes, there is a risk that the uPort IDs may be compromised if an attacker developed a malicious application that would obtain full access to the registry.

Sovrin [35] is a decentralized identity network implemented using permissioned HyperLedger to employ digital credentials that can be used off-chain. With Sovrin, users can create multiple separate identifiers with distinct asymmetric key pair and adopts zero-knowledge proof to protect users' privacy. Sovrin enables off-chain validation of proofs with third-party organizations directly in a secure channel. In Sovrin, users must rely on organizations acting on their behalf, also called "steward" nodes to maintain the distributed ledger. Users' credentials will be distributed and replicated among the steward nodes so that users' key pairs are recoverable. However, the Sovrin network continues to face several challenges, such as verifying the data recipient, ecological construction of on-chain information, and some credential information that might be leaked to third parties depending on the steward nodes' policies, which makes adoption difficult in the highly regulated

Table 1
Comparison of related work on the major identity management features.

Features	ShoCard	Sovrin	uPort	MediLinker	Our approach
User consent and identity control	Control of creation and disclosure	Users have to rely on organizations to have control of their identity	Users have control but identity attributes are stored in the registry	Full control of identity through HyperLedger Indy's identity management	Full control of identity through NFT authentication
Centralized authority	Centralized server validation	No centralized authority with a web of trust	Centralized registry and trusted attribute providers	No centralized authority	No centralized authority
Sensitive data/medical records management and sharing	Encrypted credentials are stored in a centralized server	Credentials can be stored on mobile but potentially leaked to agents	Identity attributes are stored in centralized registry	Store on personal devices, data exchange protocol is unclear	Stored and exchange through distributed file system after encryption
Scalability	Highly scalable through a centralized server	High scalability through rings of nodes	Off-chain storage and processing, still challenging	Scalability is not tested	Raft consensus/blockchain adapter buffer
Selective disclosure	Not applicable	Zero-knowledge-based selective disclosure	Users can selectively disclose or change their attributes	Supported in the introduction but implementation is not clear	Patients can decide what information are sharable

healthcare industry.

MediLinker [36] is a user-centric identity and consent management system using HyperLedger. Patients use private keys stored on their mobile devices to prove their identity and share their medical information with healthcare providers with their consent. A 30-participant, multi-site, real-world pilot study is conducted to evaluate the usability, functionality, and feasibility of MediLinker. The assessment offers proof of the possible use of a patient-centric, blockchain-based identity management system in healthcare settings. However, details of technical designs are lacking, and several challenges would need to be addressed in their future research revolving around scalability testing, security settings, data exchange protocol, and users' key management.

In this work, we have designed and implemented a blockchain architecture for the SSI-empowered NFT's overall management structure. With NFT-based patient tokenization, patients can own and control their identities, allowing them to decide how and with whom their PHI and medical history are shared, leading to a transparent and efficient data-sharing system that prevents data breaches caused by unauthorized access. We have optimized the existing approaches for identity management and tailored the design to healthcare settings. Our proposed architecture, specifically, generates a separate NFT only for each patient's identity management, contains unique settings to safeguard on-chain and off-chain data security through rigorous encryption and hashing mechanisms, ensures the legitimacy of each participant by conducting an on-site identity check, and permits selective disclosure so that patients can decide which parts of their medical records can be shared. We have also conducted case studies and scalability tests to assess the feasibility and performance of the architecture, as described in the remaining sections.

3. Methods

3.1. Environment setup

In this work, we have adopted the Quorum blockchain, developed by J.P Morgan, for the system design and implementation. In comparison to the original blockchain, the Quorum blockchain uses a Raft-based consensus mechanism, which produces an improvement in scalability and efficiency [37,38]. All participating sites need to provide a blockchain node, which is an electronic device that runs blockchain services and obtains permission from the blockchain originator before joining. There are three types of nodes in our network, as shown in Fig. 1: (1) the *full node*, which stores the latest state of the blockchain, validates and sends transactions, and is able to retrieve the whole blockchain history by tracing transactions stored in the blocks; (2) the *archive node*, which contains all functions of the full node and stores the entire blockchain's data from the origination, and (3) the *light node*, which can validate and send transactions, but only stores part of the block's information to reduce the requirements on its hardware specifications. It is noteworthy to mention that the archive node, which is not centralized [39], just stores all blockchain transactions to allow for quick indexing, which is crucial for boosting the effectiveness of the auditing process. To participate in the blockchain system, the originator or authority must provide an archive node which can be an electronic device with sizable storage capacity to initiate the blockchain and formulate the rules through smart contracts. This allows the authority to have a fast audit process by tracing the transactions locally. Each healthcare facility is required to provide a full node which can be a desktop installing a Quorum blockchain client to connect on-chain and off-chain activities, and an InterPlanetary File System (IPFS), which is an innovative, peer-to-peer distributed file system using cryptographic hash for file storing and retrieving that has been proved efficient, stable, and secure

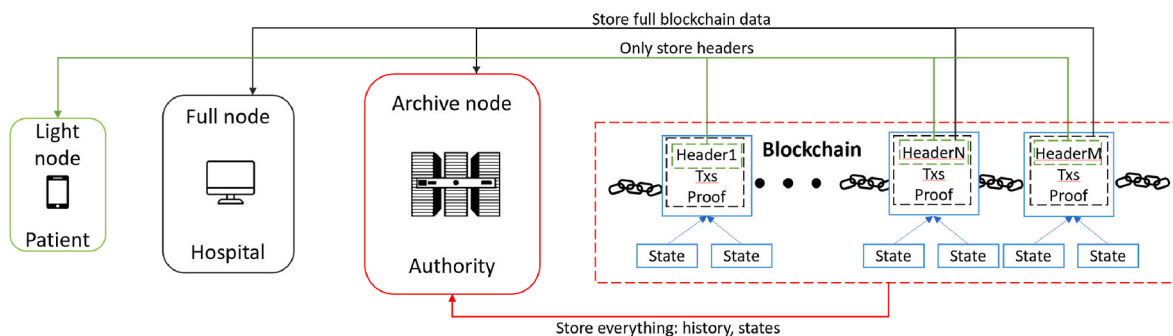


Fig. 1. Three types of nodes in the blockchain system: light nodes that only store blockchain headers, full nodes that store all blockchain data, and archive nodes that include all blockchain information to support a rapid index.

[40]. A Remote Procedure Call (RPC) server is built inside the full node to engage with off-chain activities, such as sending query requests to the EHR database protected by the firewall. Full nodes and archive nodes must follow the local Health IT regulations. The patient needs to provide a light node which can be their smart mobile device to send transactions, authenticate NFT requests, audit the NFT access, and keep a record on both the blockchain and their local device.

To begin, the authority (i.e., the Department of Health and Human Services) uses the archive node to deploy the genesis block (which holds the blockchain’s specifications) as well as the smart contracts, which specify the system’s rules. Each healthcare facility will join the blockchain system by obtaining certain permissions from the authority, such as sending the authority identity proof, after which their full nodes will be added to the blockchain system. Users will receive their tokens as soon as they register in the system. To ease the complex installation procedure, the whole installation and setup process is implemented as an executable application or mobile app.

3.2. System design

The overall architecture, as shown in Fig. 2, contains four modules: (1) Creation Module, which permits users to create tokens through blockchain, (2) Linkage Module, which links the patient account from each health system to the patient token, (3) Authentication Module, which allows patients to permit healthcare providers to access their token to retrieve their patient’s medical history from multiple healthcare facilities, and (4) Exchange Module, which involves validation of the legitimacy of the token and the HIE process.

First, patients must download and install the NFT app on their mobile devices. The app will then convert the mobile device into a light node and automatically create a blockchain account for the patient. To receive their NFT, patients only need to set up their biometric information and link to their blockchain account. Each healthcare facility and authority must install the NFT software in order to convert their electronic device into a full node and an archive node, respectively. Once patients have proven their identities and ownership of the NFT in person to the healthcare facility from the prior visits, the administrator needs to extract the local patient identifiers (in accordance with local Health IT

regulations) and input them into the NFT through a Graphical User Interface (GUI) in the software. On the mobile app, patients may customize the sharing permissions for each record. Records that do not have sharing permission will be automatically blocked from further access. When patients want to share their medical history with healthcare providers they need SSI to validate their identity by verifying their biometric information using the mobile app. Patients can grant NFT permission to the healthcare provider by scanning a QR code or NFC tap to add the provider’s blockchain account to their allow list in the NFT; then, based on the information stored on the NFT, HIE will proceed automatically between the remote and local healthcare facilities. The software deployed on their facility’s full node will allow healthcare providers to view the permitted medical history.

The technical details of backend activities are described in the following sections: Creation module outlined transactions from the light node to the blockchain, specifically the smart contract for NFT creation, following Linkage module presented the operations on full nodes to extract local patient identifiers, link with NFT, and create local index table, then the Authentication module explained the process on the light node for patient authentication on their records and the ownership of NFT, finally, the Exchange module discussed the actions of the data request, data exchange, data encryption, and decryption occurred during the HIE process.

3.2.1. Creation module

A blockchain account (a unique pair of public-private keys) will be generated for each user once each blockchain node joins the system, such as one patient per light node, system administrators from their full nodes, and the archive node. When a patient’s account is ready, the light node will automatically send a transaction to the blockchain to create an NFT (a unique smart contract only for the patient to use). The logic of the smart contract design and part of the source code is shown in Fig. 3. Patients need to set up face, voice, or fingerprint recognition through the app to link with their private keys. The NFT contains the following visible information: (1) the patient’s blockchain account address (derived from the public key) and (2) hospital blockchain accounts from the patient’s past visits and the following confidential information: (1) hashes of biometric information linked to the patient’s private key, (2)

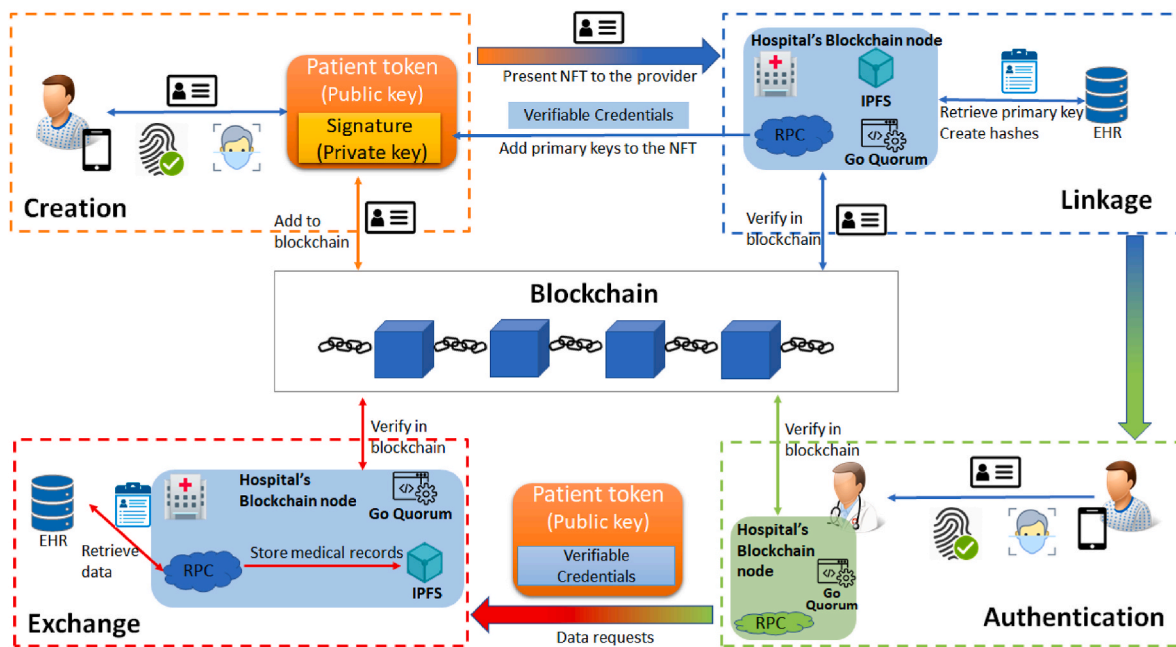


Fig. 2. Four modules contained in the overall architecture for the use of NFT for HIE authentication, beginning with the creation of NFTs through blockchain, followed by the linkage of remote patient IDs across healthcare facilities, and ending with patient authentication for ownership of NFTs and permission for healthcare providers to retrieve their medical history in the final Exchange module.

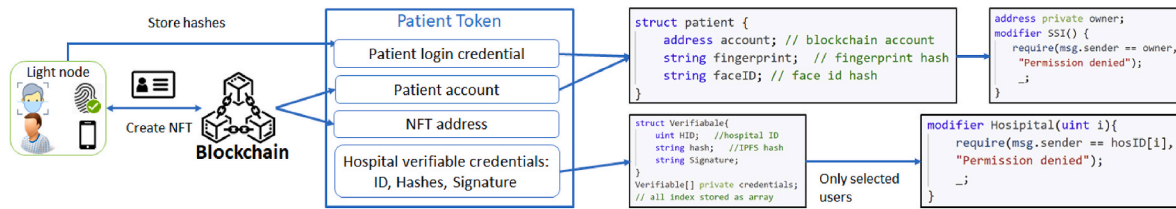


Fig. 3. The NFT structure contains patients’ login information for future validation, as well as hospital verifiable credentials, which include the healthcare facility’s signatures and hashes derived from the primary key queried from the EHR database.

signatures of the hospital who input the medical records, and (3) identifiers for the patient as well as for visits from each hospital. Several predefined modifiers prevent individuals with certain responsibilities from executing specific actions, such as enabling only patients to input their own login credentials and only healthcare facilities to enter local identifiers into the NFT.

In this module, patients can receive their NFTs by creating an exclusive smart contract through their mobile devices. The biometric data is stored locally, and the hashes are kept inside the NFTs to maintain consistency so that only the patients themselves can provide proof of the ownership of the token to achieve SSI. The NFTs, however, cannot be retrieved if the patient’s mobile device is lost. They must recreate the NFT following the same procedures mentioned previously. The NFTs are made publicly available over the blockchain network, allowing all system users to consent to the authenticity of the NFTs without the need for an intermediary to reveal and validate the genuine information contained inside.

3.2.2. Linkage module

The NFTs are designed to be linked with patient accounts across healthcare facilities (shown in Fig. 4), as the primary goal is to unify patient identification. We established the following assumptions to achieve the linkage: each healthcare facility must set up a blockchain node in accordance with its local Health IT regulations to store parts of the EHR, such as patient IDs and visit IDs, inside the blockchain node. To provide another layer of security for the NFT, the healthcare facility requires a two-step verification once the patient shows their NFT, including SMS verification and the presentation of a physical ID card. Next, the administrator from each healthcare facility needs to perform the following steps to store the information in the NFTs: (1) access their database to extract the primary keys (the attributes that can uniquely identify particular instances of the patient), (2) store the primary keys, the hashes of the primary keys, and the NFT address in the index table created on the blockchain node, and (3) send the hashes to the NFT’s smart contract address through blockchain node. When the healthcare facility adds the information to the NFT, a signature (produced using the Diffie–Hellman key exchange protocol [41]) will be added along with each record to the smart contract. The blockchain system will execute the key generation process automatically. The private key of the healthcare facility and the public key of the patient are used to create the signature and the system will produce a verification key using the public key of the healthcare facility and the patient’s private key so that

patients may authenticate in the future. Using the Diffie–Hellman protocol, users can establish a commonsense approach to information security without exposing their own private keys while recognizing the legitimacy of the decoders’ identities.

In this module, the NFTs will be linked to all patient IDs across healthcare facilities to achieve the goal of patient tokenization. Blockchain can assure the immutability of data, anonymize patients and healthcare facilities, and provide a potential solution to the HIE process’s persistent patient matching challenge. An innovative master patient index table can be built on the blockchain node using the Diffie–Hellman encryption protocol to build quick and accurate data retrieval.

3.2.3. Authentication module

Once patients’ tokens have been fully connected across prior healthcare facilities, patients can choose what data to share, when to share it, and with whom it will be shared. This module has two key aspects: (1) the selective sharing of their records and (2) SSI verification. Patients can set a flag to the records that they don’t want to exchange in the NFT to block future access through the smart contract. Those records still exist in the NFTs for personal use but cannot be accessed by others. When patients share their NFT with healthcare providers, those records will not appear in the patient’s medical history. Patients have the flexibility to unflag such records whenever they are willing to share them. In previous research, we investigated a patient-centric HIE that would allow patients to control which parts of their medical history could be shared [11]. Doing this requires a more complex encryption mechanism on both the remote blockchain node and the recipient’s blockchain node, beyond the scope of this work.

To achieve SSI, patients first must validate themselves by authenticating their biometric information on their mobile devices to log in to the blockchain app. The NFT will then be verified by matching the private keys linked to the biometric data with the public key stored on the blockchain. The patient has proven their ownership of the token by logging in successfully. The patient then must provide the NFT to the recipient by adding their blockchain account to the patient’s NFT smart contract’s allowed list. After 10 min, the permission will be revoked. Instead of manually inputting the account, this step can be accomplished by scanning healthcare providers’ QR codes or by one of several other simple alternatives.

In this module, patients are able to authenticate themselves using blockchain which ensures the next part of the process without the

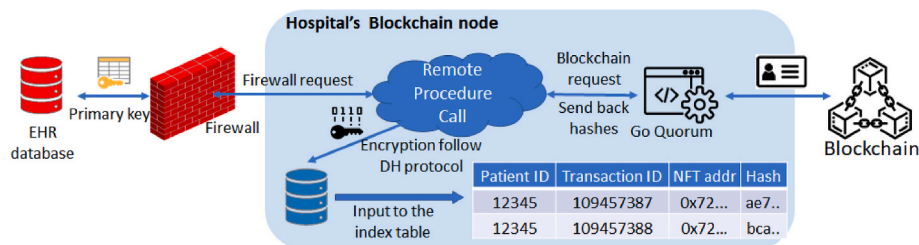


Fig. 4. After validating the authenticity of NFT using blockchain, healthcare facilities establish the linkage as follows: Obtain primary keys from the EHR database, encrypt the keys using the DH protocol, create an index table locally, and deliver the hashes to the NFT.

involvement of intermediary parties: (1) the validity of the holder’s identity, (2) the legitimacy of record issuers, and (3) the consistency of the information contained within the NFT.

3.2.4. Exchange module

The last module is the exchange module wherein HIE takes place. Patients have verified their identities using SSI-enabled NFTs after completing the preceding modules. The recipients, mostly healthcare providers, will then make requests to the remote healthcare facilities specified in the patient’s medical history contained in the NFT to acquire the patient’s health information. Since the patient has already granted permission to the recipient, the patient’s blockchain node will generate verification keys using the patient’s private key along with the public keys of remote healthcare facilities and then embed them in the NFT. The remote healthcare facility will retrieve the index keys, (the primary keys used for querying records from the EHR database) through the blockchain node after receiving the verification keys, indicating that these queries have been previously authorized by the patient.

After the records from the EHR database have been queried, they will be sent to the local blockchain node and encrypted with a random key. The encrypted data will be stored on the blockchain node’s IPFS, which will return a unique hash string for indexing. The blockchain node will then use a signature to encrypt the decryption key and hash string. Finally, the local blockchain node will send the encrypted decryption key as well as the hash string to the requester through a private blockchain transaction that involves a second layer of encryption and decryption using the sender’s and receiver’s key pairs, preventing others in the blockchain network from decoding the real text in the private transaction even though the encrypted data stored inside the IPFS cannot be decrypted without the verification keys.

When the request blockchain account receives the private transaction, the blockchain decodes the internal information and uses the verification key stored in the NFT to decipher the decryption key automatically. Next, the blockchain node will use the hash string to retrieve the encrypted data from the IPFS and then use the decryption key to decrypt the patient’s medical history. At this point, SSI and HIE are both completed.

4. Case studies and results

To demonstrate the use of our proposed system, we conducted a case study using the following scenario. A patient with chronic kidney disease lived in City A for the past ten years and future care in the new location is needed. His medical records are mostly stored at two hospitals in City A. The patient is relocated to City B due to a job transfer. As his new employment entails frequent business travel, he wants his future doctors (to be located in City B), as well as other locations, to have a better understanding of his medical history to provide him with better healthcare. However, due to concerns about provider bias, he does not want new healthcare providers to know about his history of alcohol addiction prior to his diagnosis of chronic kidney disease. He has been sobered for several years, with assurances from his primary care physician that his medical history will have no impact on his future health.

In this instance, the patient can benefit from our proposed blockchain system without worrying about identity loss due to frequent HIE. The implemented system contains five physical nodes representing one archive node for the initiator, two hospitals’ full nodes from City A, a hospital full node from City B, and a light node for the patient in the system. Each blockchain node is equipped with an Ubuntu operating system, a quorum blockchain client, and an IPFS file system. We have also used synthetic data to simulate the whole process. To originate the blockchain system, the initiator node needs to deploy a unique genesis file that specifies the use of the raft consensus mechanism, comprising the lowest difficulty level (to promote efficiency) and sufficient initial balance for each account to make transactions.

After the processes described in the creation and linkage modules are completed, the patient’s NFT becomes a unified identifier across healthcare facilities. We have built a sample GUI for this patient’s NFT, as shown in Fig. 5. All blockchain users can see the patient’s blockchain account and the NFT address. In the NFT, records are listed with index keys and authorization options. The patient has blocked access to the records related to his alcoholism. When the doctor’s blockchain account is entered into the patient’s NFT, it shows the patient has authorized access to the NFT for the doctor. Instead of manually entering the doctor’s account, as previously stated, this procedure can be simplified. Next, the doctor will receive the NFT with all sharable information. Data requests will be transmitted automatically to the blockchain accounts of the remote healthcare facilities, and the remote blockchain nodes will perform the exchange module processes. A sample data encryption and exchange process from the backend is shown in Fig. 6. We use SHA1 as a hash algorithm and OpenSSL as an encryption implementation. These methods are only for simulation purposes, but they may be customized on the blockchain node at any time. The retrieval process is the inverse of the sender’s process, as it involves retrieving encrypted the IPFS hash and decrypt key from the blockchain, decrypting the IPFS hash, retrieving encrypted data using the hash from IPFS, and decrypting the original data using verification keys generated by the patient’s blockchain node.

As the aim of this work is to use blockchain for identity validation and transfer, we have created and run the script containing the following steps every second continuously for a week: (1) the patient granting permission to a doctor’s account, (2) the doctor’s account sending data requests to the healthcare facility in City A, and (3) the healthcare facilities in City A sending the IPFS hashes to the doctor’s account. There are over 3 million transactions (authentication, data request, and sending IPFS hash values) that have been completed, with a success rate of 100% and a blockchain validation and written time of 1.17 s on average.

To further investigate the performance, especially the stability of the proposed system, and only for the simulation purpose, another test node was added to the system with 100 patient accounts. We have created a script to set up the preconditions as creating NFTs for each patient, linking the existing records to the patients’ NFT, adding the patients’ accounts to the index tables on each healthcare facility’s node, generating an IPFS hash value, and adding the provider’s account to the allow list of each NFT. We have excluded the off-chain activities for this stability test, such as the encryption and decryption process, data storing and retrieving from the IPFS, and patient identifiers extraction from the EHR databases. The following scripts were then written and executed

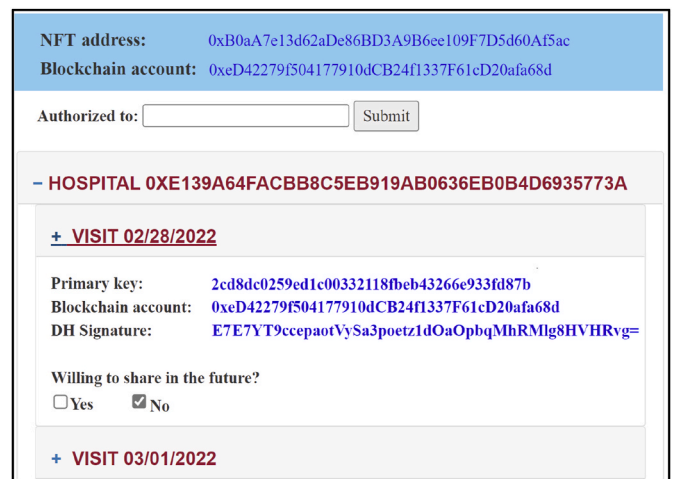


Fig. 5. GUI of NFT including the patient’s blockchain account, NFT smart contract address, and a list of previous visits’ linkage information.

```

$ tail patient773A_request01.csv
PID,TID,Primary_dianosis,Date,Discharge
1,10001,N18. 9,2282022,2282022
1,10002,N18. 9,3012022,3012022

```

Queried records

```

$ openssl rsautl -in patient773A_request01.csv -out patient773A_request01.enc -pubin -inkey random-pub.pem -encrypt
$ tail patient773A_request01.
patient773A_request01.csv patient773A_request01.enc
$ tail patient773A_request01.enc
00000R0s.p0000%0v'9.R0
00,00A()00<0K>(T000|T00x0
\00X 00!U0T0'0%E0000
0}
0a00G{\-0a30Bd000000"0U0s00J0>0000v'0c03000M0z^s0`i0,n410v010c0I[0#0y00k000@B300L]0J00A000A0.%90
00p(00n000
I000|2:00700000000I0000;4-0ym0D1$000^0T0c`$p00'0J10e0U00000.rdk0A#000p0000g9z.0]0"000ZI.n000%.n00R=00(00000w1040000000T00.|00K=L09&v
$ ipfs add patient773A_request01.enc
added QmctEFiPy7iT2RY5iSmXTPtbHDo5HbKK9Lh27tUpJXYWYP patient773A_request01.enc
384 B / 384 B [=====] 100.00%
$ echo QmctEFiPy7iT2RY5iSmXTPtbHDo5HbKK9Lh27tUpJXYWYP > ipfs_request01.txt
$ openssl rsautl -in ipfs_request01.txt -out ipfs_request01.enc -pubin -inkey public-key.pem -encrypt
$ openssl rsautl -in random-pri.pem -out random-pri.dec -pubin -inkey public-key.pem -encrypt
$ head patient.nft ipfs_request01.enc random-pri.dec > parameters.txt
$ node quorum.js parameters.txt
Pass NFT address encrypted IPFS hash and decryption keys to the smart contract

```

Encrypt records
Store encrypted records to IPFS
Encrypt hash and decryption key

Fig. 6. The backend of the encryption and exchange process occurred in the remote healthcare facility’s blockchain node.

continuously every 2 s for an hour: (1) the healthcare provider’s full node sends a transaction to each remote healthcare facility’s node to request each patient’s records, (2) after 1 s, each remote healthcare facility’s node sends a transaction to the provider’s blockchain account containing the generated IPFS hash string for each patient. This creates 200 transactions per second for the simulation. The average transaction processing time is 1.42 s for the first hour and Fig. 7 presents the system stability, which refers to the speed of block generation.

5. Discussion and conclusion

We have proposed an SSI-enabled patient tokenization system that enables patients to authenticate themselves without the need for an intermediary party to minimize the risk of PHI breaches. By utilizing blockchain’s unique features, healthcare facilities may anonymously link the local patients’ identifiers to the NFT for future HIE requests. Smart contracts and security properties help to ensure that the processes comply with the regulations. Patients can always audit the records’ access history by tracking the blockchain history. In our case study, we mimicked the NFT authentication and exchange processes to assess the system’s feasibility. Compared to the case study, the stability test has a higher transaction throughput with a similar transaction processing time. We empirically confirmed the viability of the proposed blockchain system through the success rate and stability.

We have carried out a cost analysis from the following three perspectives [42] to further assess the possibility of adopting the proposed blockchain system for patient identity authentication: (1) implementation costs which include software installation and hardware investment; (2) switching costs based on market share and clients preferences; and (3) anticipated losses if the blockchain system fails. Firstly, since the

blockchain system is an open-source distributed ledger technology, all users can join and register an NFT without extra cost as long as their device can run mobile Apps or computer programs. No particular hardware is necessary to deploy the proposed system. Clinical sites can keep their computers as full nodes, while patients can continue to use their mobile devices as light nodes. The authority is encouraged to provide enough storage for archive nodes, but it is not a requirement. Secondly, the proposed blockchain system is not intended to replace healthcare information systems (HISs); rather, it is meant to address the persistent challenge of PHI breaches and patient matching. The blockchain system relies on current local patient IDs to give patients control over their identities across clinical sites through unique NFTs. By implementing a Remote Procedure Call on blockchain nodes, the proposed blockchain system is interoperable and can be implemented as an add-on component to the existing HISs. There is no need to switch or replace the current HISs. Finally, blockchain technology has the feature of durability and robustness so that it has resistance to a single point of failure. Only when all blockchain nodes fail does the blockchain system terminates, hence we will deploy a blockchain node to mitigate this risk. When blockchain nodes fail and rejoin the network, they immediately synchronize previously completed transactions from other blockchain nodes. In summary, we analytically infer that the proposed blockchain is cost-effective to be adopted for patient identity management tasks.

However, several limitations exist in this architecture. Patients must present their proof of identity in person for the healthcare facilities to verify their identity authenticity. Their physical presence is necessary for the identifier linkage because healthcare data is subject to strict rules. Alternative approaches, such as signing consent paperwork with a digital signature, may help to alleviate the travel-related burden, but they require numerous sites, such as healthcare facilities, insurance, and

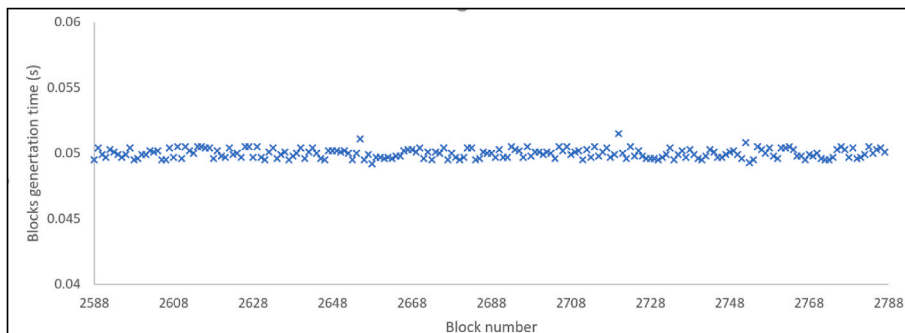


Fig. 7. The generation time consumption of 200 randomly selected continuous blocks from the proposed system’s stability test results.

government agreements [43]. Another limitation is related to the HIE. Through the NFT, patients can control their identities but they cannot fully control the records which have been exchanged. Our prior work developed a comprehensive design for the blockchain architecture of patient-centric HIE, allowing patients to personalize as well as segment their records thus determining which parts of their health records can be shared. This approach can also be used in conjunction with NFT, i.e., by storing all medical records on the mobile device and linking with the NFT so that patients have complete control over what data they will share when they want to share it, and with whom their data is shared. An additional limitation is the scalability constraint of the blockchain system. Based on our blockchain research experience, the number of nodes barely affected the performance of the blockchain system. There may be a slight difference in milliseconds due to propagating the transactions to more blockchain nodes. However, processed transactions per second are the main limitation of all blockchain systems. From the stability test result, our proposed system can effectively handle 200 transactions per second. According to an incomplete statistical observation of 860.4 million ambulatory visits in 2018 [44], or around 27 patients per second in the United States, our approach should handle the patients' identification verification demands. Furthermore, our prior research has developed blockchain adapters to buffer unprocessed transactions to keep the blockchain scalable, stable, and efficient. However, many factors can affect the frequency of patient visits, such as disease outbreaks. Still, there are other challenges such as mobile device specs, internet transmission speed limits, and, most importantly, government and healthcare facility restrictions.

In future work, we will continue to assess the blockchain protocol and formulate strategies to fundamentally address the blockchain's scalability issue. We will also keep exploring the potential of NFT in important data exchanges, such as integrating NFTs with the cryptocurrency concept to establish an incentive mechanism for data sharing. We will also provide a comprehensive functionality design for HIE, such as personalized data segmentation, partial EHR indexing, and smart contracts for regulation enforcement, based on NFT authentication. Large-scale simulations using real-world data will be conducted to further evaluate the feasibility, stability, scalability, and security of the system.

Declaration of competing interest

None Declared.

Acknowledgments

This research was supported in part by grant 72125009 from the National Natural Science Foundation of China, grant 2022YFF1203000 from the National Key R&D Program of China, 62102008 from the National Natural Science Foundation of China, 2020BD004 from PKU-Baidu Fund, and the Shumaker Endowment in Biomedical Informatics.

References

- V. Liu, M.A. Musen, T. Chou, Data breaches of protected health information in the United States, *JAMA* 313 (14) (2015) 1471–1473.
- A.H. Seh, M. Zarour, M. Alenezi, A.K. Sarkar, A. Agrawal, R. Kumar, R. Ahmad Khan, *Healthcare Data Breach: Insight. Implicate*. 8 (2) (2020) 133.
- A. McLeod, D. Dolezel, Cyber-analytics: modeling factors associated with healthcare data breaches, *Decis. Support Syst.* 108 (2018) 57–68.
- Portal HHS. Breach, Notice to the Secretary of HHS breach of unsecured protected health information affecting 500 or more individuals. United States Department of Health and Human Services (HHS), in: Book Notice to the Secretary of HHS Breach of Unsecured Protected Health Information Affecting 500 or More Individuals. United States Department of Health and Human Services (HHS), edn., 2022.
- T.M. Keane, C. O'Donovan, J.A. Vizcaino, The growing need for controlled data access models in clinical proteomics and metabolomics, *Nat. Commun.* 12 (1) (2021) 5787.
- J.C. Mandel, J. Pollak, K.D.J.o.M.I.R. Mandl, 11, in: *The Patient Role in a Federal National-Scale Health Information Exchange*, 24, 2022, e41750.
- B.H. Just, D. Marc, M. Munns, R. Sandefer, Why patient matching is a challenge: research on master patient index (MPI) data discrepancies in key identifying fields, *Spring. Perspect. Health Inf. Manag.* 13 (2016) 1e, 1e.
- G.J. Kuperman, Health-information exchange: why are we doing it, and what are we doing? *J. Am. Med. Inf. Assoc.* 18 (5) (2011) 678–682.
- P. Esmailzadeh, Identification of barriers affecting the use of health information exchange (HIE) in clinicians' practices: an empirical study in the United States, *Technol. Soc.* 70 (2022), 102007.
- P.T.S. Liu, Medical record system using blockchain, big data and tokenization, in: *Book Medical Record System Using Blockchain, Big Data and Tokenization*, edn., Springer International Publishing, 2016, pp. 254–261.
- Y. Zhuang, L.R. Sheets, Y.W. Chen, Z.Y. Shae, J.J.P. Tsai, C.R. Shyu, A patient-centric health information exchange framework using blockchain technology, *IEEE J. Biomed. Health Informatic.* 24 (8) (2020) 2169–2176.
- T.-T. Kuo, H.-E. Kim, L. Ohno-Machado, Blockchain distributed ledger technologies for biomedical and health care applications, *J. Am. Med. Inf. Assoc.* 24 (6) (2017) 1211–1220.
- I. Yaqoob, K. Salah, R. Jayaraman, Y. Al-Hammadi, Blockchain for healthcare data management: opportunities, challenges, and future recommendations, *Neural Comput. Appl.* 34 (14) (2022) 11475–11490.
- P. Zhang, J. White, D.C. Schmidt, G. Lenz, S.T. Rosenbloom, FHIRChain: applying blockchain to securely and scalably share clinical data, *Comput. Struct. Biotechnol. J.* 16 (2018) 267–278.
- Y. Zhuang, L. Zhang, X. Gao, Z.-Y. Shae, J.J.P. Tsai, P. Li, C.-R. Shyu, Re-Engineering a clinical trial management system using blockchain technology: system design, development, and case studies, *J. Med. Internet Res.* 24 (6) (2022), e36774.
- L. Hang, C. Chen, L. Zhang, J. Yang, Blockchain for applications of clinical trials: Taxonomy, challenges, and future directions, *IET Commun.* 16 (2022) 2371–2393, <https://doi.org/10.1049/cmu2.12488>.
- P. Syllim, F. Liu, A. Marcelo, P. Fontelo, Blockchain technology for detecting falsified and substandard drugs in distribution: pharmaceutical supply chain intervention, *JMIR Res. Protoc.* 7 (9) (2018) e10163, e10163.
- T.-T. Kuo, A. Pham, Quorum-based model learning on a blockchain hierarchical clinical research network using smart contracts, *Int. J. Med. Inf.* (2022), 104924.
- Y. Zhuang, L. Sheets, Z. Shae, J.J.P. Tsai, C.-R. Shyu, Applying blockchain technology for health information exchange and persistent monitoring for clinical trials, *AMIA Annu. Symp. Proc.* 2018 (2018) 1167–1175.
- Q. Wang, R. Li, Q. Wang, S. Chen, Non-fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges, 2021 arXiv preprint arXiv:2105.07447.
- M. Nadini, L. Alessandretti, F. Di Giacinto, M. Martino, L.M. Aiello, A. Baronchelli, Mapping the NFT revolution: market trends, trade networks, and visual features, *Sci. Rep.* 11 (1) (2021) 1–11.
- P. Zhang, T.-T. Kuo, The feasibility and significance of employing blockchain-based identity solutions in health care, in: S. Patnaik, T.-S. Wang, T. Shen, S.K. Panigrahi (Eds.), *Blockchain Technology and Innovations in Business Processes*, Springer Singapore, 2021, pp. 189–208.
- K. Kostick-Quenet, K.D. Mandl, T. Minssen, I.G. Cohen, U. Gasser, I. Kohane, A. L. McGuire, How NFTs could transform health information exchange, *Science* 375 (6580) (2022) 500–502.
- L. Lu, J. Han, Y. Liu, L. Hu, J.-P. Huai, L. Ni, J. Ma, Pseudo trust: zero-knowledge authentication in anonymous P2Ps, *IEEE Trans. Parallel Distr. Syst.* 19 (10) (2008) 1325–1337.
- T.K. Mackey, A.J. Calac, B.S. Chenna Keshava, J. Yracheta, K.S. Tsosie, K. Fox, Establishing a blockchain-enabled Indigenous data sovereignty framework for genomic data, *Cell* 185 (15) (2022) 2626–2631.
- A. Mühle, A. Grüner, T. Gayvoronskaya, C. Meinel, A survey on essential components of a self-sovereign identity, *Comput. Sci. Rev.* 30 (2018) 80–86.
- A. Tobin, D. Reed, *The Inevitable Rise of Self-Sovereign Identity*, vol. 29, The Sovrin Foundation, 2016, 2016.
- X. Sun, F.R. Yu, P. Zhang, Z. Sun, W. Xie, X. Peng, A survey on zero-knowledge proof in blockchain, *IEEE Network* 35 (4) (2021) 198–205.
- M. Davie, D. Gisolfi, D. Hardman, J. Jordan, D. O'Donnell, D. Reed, The trust over ip stack, *IEEE Commun. Standard Magazine.* 3 (4) (2019) 46–51.
- W.J. Gordon, C. Catalini, Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability, *Comput. Struct. Biotechnol. J.* 16 (2018) 224–230.
- Y. Liu, D. He, M.S. Obaidat, N. Kumar, M.K. Khan, K.-K. Raymond Choo, Blockchain-based identity management systems: A review, *J. Netw. Comput. Appl.* 166 (2020) 102731.
- Q. Feng, D. He, S. Zeadally, M.K. Khan, N.J.J.o.N. Kumar, C. Applications, in: *A survey on privacy protection in blockchain system*, 126, 2019, pp. 45–58.
- S. ShoCard, Travel identity of the future—white paper, in: *Book Travel Identity of the Future—White Paper*, edn., May, 2016.
- C. Lundkvist, R. Heck, J. Torstensson, Z. Mitton, M.J.U.h.w.u.m.u.w.D.p. Sena, Uport: A Platform for Self-Sovereign Identity, 2017.
- A. Tobin, D.J.T.S.F. Reed, 2016, in: *The inevitable rise of self-sovereign identity*, 29, 2016, p. 18.
- A. Khurshid, C. Holan, C. Cowley, J. Alexander, D.T. Harrell, M. Usman, I. Desai, J. R. Bautista, E. Meyer, Designing and testing a blockchain application for patient identity management in healthcare, *JAMIA Open* 4 (3) (2021) o0aa073.
- A. Baliga, I. Subhod, P. Kamat, S. Chatterjee, Performance Evaluation of the Quorum Blockchain Platform, 2018 arXiv preprint arXiv:1809.03421.
- C. Cachin, M. Vukolić, *Blockchain Consensus Protocols in the Wild*, 2017 arXiv preprint arXiv:1707.01873.

- [39] J.-Y. Kim, J. Lee, Y. Koo, S. Park, S.-M. Moon, Ethanos: efficient bootstrapping for full nodes on account-based blockchain, in: *Book Ethanos: Efficient Bootstrapping for Full Nodes on Account-Based Blockchain*, edn., 2021, pp. 99–113.
- [40] Y. Chen, H. Li, K. Li, J. Zhang, An improved P2P file system scheme based on IPFS and Blockchain, in: *2017 IEEE International Conference on Big Data (Big Data)*, Boston, MA, USA, 2017, pp. 2652–2657, <https://doi.org/10.1109/BigData.2017.8258226>.
- [41] S. Michael, G. Tsudik, M. Waidner, Diffie-Hellman key distribution extended to group communication, in: *Proceedings of the 3rd ACM Conference on Computer and Communications Security*, 1996, pp. 31–37.
- [42] K. Altinkemer, T. Wang, Cost and benefit analysis of authentication systems, *Decis. Support Syst.* 51 (3) (2011) 394–404.
- [43] P. Esmailzadeh, Benefits and concerns associated with blockchain-based health information exchange (HIE): a qualitative study from physicians' perspectives, *BMC Med. Inf. Decis. Making* 22 (1) (2022) 80.
- [44] 'National Ambulatory Medical Care Survey, National summary tables, in: *Book National Ambulatory Medical Care Survey: 2018 National Summary Tables*, edn., 2018.