A Blockchain-Based Personal Health Knowledge Graph for Secure Integrated Health Data Management

Juan Li
Department of Computer Science
North Dakota State University
Fargo, USA
j.li@ndsu.edu

Vikram Pandey
Department of Computer Science
North Dakota State University
Fargo, USA
vikram.pandey@ndsu.edu

Rasha Hendawi
Department of Computer Science
North Dakota State University
Fargo, USA
rasha.hendawi@ndsu.edu

***Abstract*— The increasing use of electronic health records (EHRs) and wearable devices has led to the creation of massive amounts of personal health data (PHD) that can be utilized for research and patient care. However, managing and integrating various types of PHD from different sources poses significant challenges, including data interoperability, data privacy, and data security. To address these challenges, this paper proposes a blockchain-based personal health knowledge graph for integrated health data management. The proposed approach utilizes knowledge graphs to structure and integrate various types of PHD, such as EHR, sensing, and insurance data, to provide a comprehensive view of an individual's health. The proposed approach utilizes blockchain to ensure data privacy and security. By storing PHD on a decentralized blockchain platform, patients have full control over their data and can grant access to specific entities as needed providing enhanced privacy and security.**

Keywords— Blockchain, Personal Health Data, Knowledge Graphs, Data Interoperability, Privacy, Security, Wearable Devices, Electronic Health Records.

I. INTRODUCTION

In recent years, there has been a significant increase in the amount of health-related data generated in our everyday lives. This data comes from a wide range of sources and includes electronic health records (EHRs), wearable devices, health-related apps, and social determinants of health. EHRs contain a wealth of information about a patient's health, including medical history, diagnoses, medications, and treatment plans. As more healthcare providers adopt electronic health record systems, the amount of EHR data generated continues to increase. Wearable devices such as smartwatches, fitness trackers, and medical devices generate a large amount of health-related data. These devices can collect data on physical activity, heart rate, sleep patterns, blood pressure, and other vital signs. With the increasing popularity of wearable devices, the amount of data generated continues to grow. Health-related apps are another source of data, providing individuals with a wide range of tools for managing their health. These apps can track diet and exercise, monitor symptoms, and provide reminders for medications and appointments. The data generated by these apps can provide valuable insights into an individual's health and can be integrated with other sources of health data. Social determinants of health, such as economic status, education, and social support, also play a significant role in an individual's overall health. These factors can be measured and analyzed to provide a comprehensive view of an individual's health and to identify potential health disparities.

The amount of health-related data generated in our everyday lives is vast and continues to grow. This data provides valuable insights into an individual's health and can be utilized for research and patient care. However, managing and integrating various types of PHD from different sources poses significant challenges, including data interoperability, data privacy, and data security.

To address these challenges, there is a need for a comprehensive and integrated approach to managing PHD. In this paper, we propose a personal health knowledge graph (PHKG), which organizes PHD in a graph format, where nodes represent entities such as patients, health practitioners, medical records, and wearable devices, and edges represent relationships between them. This approach enables the integration and analysis of various types of PHD, providing a comprehensive view of an individual's health. In addition, we propose a blockchain architecture to provide a secure and tamper-proof way of storing and sharing PHKG, while also enabling patients to have full control over their data and to share their data with trusted entities as needed. The use of smart contracts also enables the automation of access control and data sharing policies, providing enhanced privacy and security.

Overall, the use of a personal health knowledge graph and blockchain technology can provide a comprehensive and secure approach to managing PHD, enabling the integration and analysis of various types of data for research and patient care, while also providing enhanced privacy and security for patients. Evaluations using use cases and simulation have demonstrated that the proposed system is secure, scalable, and enables effective data sharing and integration.

II. RELATED WORK

Blockchain technology [1] [2] has been gaining attention in the healthcare industry as a promising solution for the secure and efficient management of personal health data. Several studies have proposed the use of blockchain for personal health record (PHR) management, data sharing, and collaborative healthcare [3]. In this section, we discuss some of the related work in this domain.

This work was supported by the National Science Foundation (NSF) with award numbers: 1722913 and 2218046.

Lee et al. [4] proposed a blockchain-based personal health record exchange system that leverages a permissioned blockchain to ensure privacy and scalability. The system allows patients to securely share their health records with healthcare providers while retaining control over their data. The authors showed that their system is more efficient and scalable compared to traditional centralized systems. Cernian et al. [5] proposed PatientDataChain that uses blockchain technology to create a decentralized healthcare infrastructure that incorporates a trust layer in the healthcare value chain. The system collects specific data from patients' medical records and integrates them into a unitary personal health records (PHR) system, where the patient is the owner of their data. Wazid et al. [6] proposed a secure communication mechanism for the exchange and storage of healthcare data using Blockchain-enabled Secure Communication Mechanism for Internet of Things-driven Personal Health Records (BIPHRS). They discussed various threats and security attacks on healthcare systems and compares the proposed BIPHRS with existing blockchain-enabled security schemes. Chen et al. [7] proposed a blockchain-enabled framework for the earlier detection of diabetes using various machine learning classification algorithms, while maintaining the EHRs of patients in a secure manner. The framework combines symptom-based disease prediction, Blockchain, and interplanetary file system, with patient health information collected via wearable sensor devices. The information is then processed by an ML model and stored in the Blockchain with the approval of the patient and practitioner.

At the same time, there has been increasing interest in the use of knowledge graphs for managing and integrating health data. Several studies have explored the use of knowledge graphs for clinical decision support and data integration[8][9][10]. For example, Shi et al.[11] studied the heterogeneous textual medical knowledge and proposed to organize and integrate the TMK into conceptual graphs. They designed a mechanism to automatically retrieve knowledge form the knowledge graph. Another study by Tal et al.[12] proposed a knowledge graph-based system for disease risk prediction. The system integrates various data sources, including electronic health records, environmental data, and social media data, to generate risk prediction models. Li et al. [13] proposed a systematic approach to construct medical knowledge graph from large scale EHRs. Their study used a big-data platform of a 3A-class hospital in China and the constructed health knowledge base contains 9 entity types, totally 22,508 entities.

In addition to clinical decision support and disease risk prediction, knowledge graphs have also been used for drug discovery and development[14][15]. The Linked Open Drug Data (LODD) cloud by Samwald et al.[16] is a knowledge graph that integrates information from various drug-related data sources, such as drug targets, pharmacological actions, and drug-drug interactions, to facilitate drug discovery and development.

These studies demonstrate the potential of knowledge graphs for managing and integrating health data. However, there is still a need for further research to explore the scalability, privacy, and security aspects of using knowledge graphs in the healthcare domain.

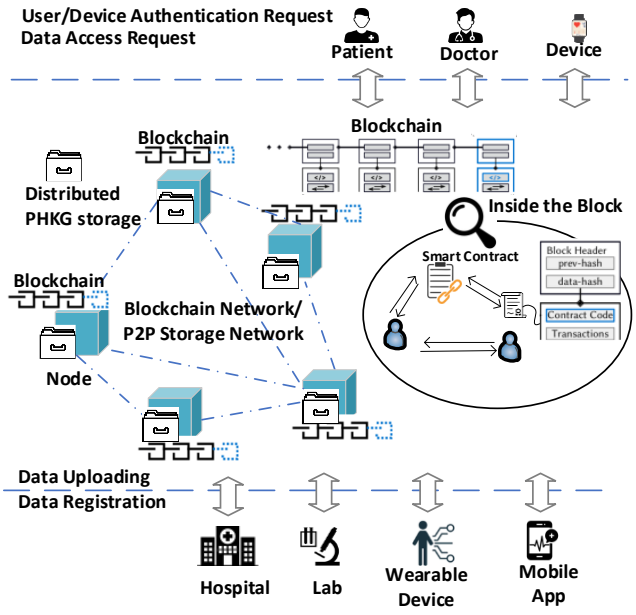

Fig. 1 System Architecture

III. System Design

The system architecture, illustrated in Fig. 1, involves various users, such as patients, doctors, health devices, and insurance agencies, who have different roles in terms of data ownership, production, and consumption. The proposed system utilizes a decentralized peer-to-peer (P2P) network that has a dual purpose: it provides a secure blockchain for system security, including authentication, authorization, and access control, and it also implements a P2P-based distributed storage to store users' personal health knowledge graph. To access the system, users must register through the blockchain's smart contract, and data access is also through the blockchain. Personal health knowledge graph data is divided into smaller parts and distributed across the P2P network. It is important to note that the personal health knowledge graph data is not stored in the blockchain itself.

A. Personal Health Knowledge Graph

The first step is to construct a high-level personal health ontology as the schema for the personal health knowledge graph. An ontology can provide a formalized and standardized way of representing and organizing the concepts and relationships within a particular domain, in this case, personal health. Constructing a high-level personal health ontology would involve defining the key concepts and relationships relevant to personal health, such as medical conditions, symptoms, treatments, medications, and healthcare providers [17][18]. This ontology could be used as a schema for the personal health knowledge graph, providing a standardized structure for organizing and representing personal health data. For example, the ontology might include a concept of "diabetes," which would have relationships to other concepts such as "blood glucose levels," "insulin therapy," and "complications." Each of these concepts could then be represented as nodes in the personal health knowledge graph, with edges connecting them to other related concepts.

We follow and adopt the HL7 FHIR [19] standard to design the ontology. Having a standardized ontology can help to ensure consistency and interoperability across different sources of personal health data. It can also provide a framework for developing intelligent applications that can reason over personal health data and provide personalized insights and recommendations. We extended an HL7 FHIR-based ontology proposed in [20]. It represents the domain entities of a personal health record. To improve semantic interoperability and knowledge sharing, ontology's entities were linked with classes from ontologies available on BioPortal. The connection was based on BioPortal's PURL (Persistent Uniform Resource Locators) identifiers to establish a semantic link to existing medical vocabularies, such as SNOMED CT or LOINC. We expanded the ontology to include other aspects of a user, including profiles, lifestyle interventions, healthcare providers, and data generated by health-related wearable devices. For example, additional classes PhysicalActivity, Diet, SmokingStatus, and AlcoholConsumption, as well as their sub-classes and properties, were added to reflect an individual's lifestyle. These classes and properties can provide additional information about an individual's lifestyle. Fig. 2. shows part of the ontology.

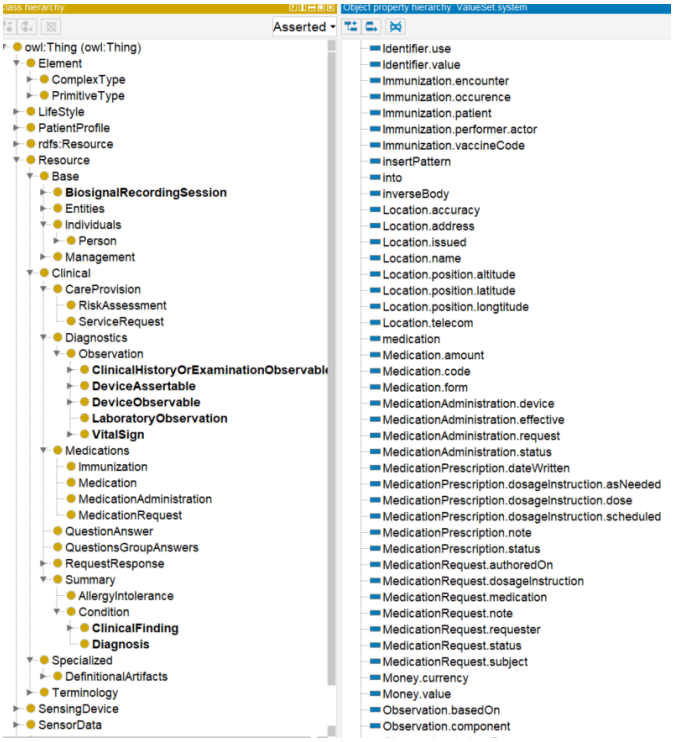

Fig. 2 A snapshot the ontology. (Partically adopted from [20])

The following step is to locate all available health data about an individual, including EHRs, medical test results, medication history, and personal health data such as symptoms, lifestyle choices, and family history, and wearable device data. These data can reside in healthcare providers, health tracking apps, and other sources. Once the data is collected, they need to be mapped with the ontology to link to their schema, which enables interoperability across different systems and applications.

Next, the integrated data is represented as a graph database, where each piece of data is a node in the graph and the relationships between the data points are represented as edges. For example, a person's medical conditions would be represented as nodes, with edges connecting them to related nodes such as medications, test results, and symptoms. Here is an example of data in a personal knowledge graph presented in triple format using RDF syntax:

<https://cs.ndsu.edu/PHKG/patient1> <http://schema.org/name> "John Smith".
<https://cs.ndsu.edu/PHKG/patient1> <http://schema.org/birthDate> "1980-01-01"^^<http://www.w3.org/2001/XMLSchema#date> .
<https://cs.ndsu.edu/PHKG/patient1> <http://schema.org/gender> "male" .
<https://cs.ndsu.edu/PHKG/patient1> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://schema.org/Person> .
<https://cs.ndsu.edu/PHKG/patient1/condition1> <http://schema.org/name> "Diabetes" .
<https://cs.ndsu.edu/PHKG/patient1/condition1> <http://schema.org/startDate> "2010-05-01"^^<http://www.w3.org/2001/XMLSchema#date> .
<https://cs.ndsu.edu/PHKG/patient1/condition1>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://cs.ndsu.edu/PHKG/Condition> .<https://cs.ndsu.edu/PHKG/patient1>
<http://cs.ndsu.edu/PHKG/hasCondition>
<https://cs.ndsu.edu/PHKG/patient1/condition1> .

<https://cs.ndsu.edu/PHKG/patient1/medication1> <http://schema.org/name> "Metformin" .
<https://cs.ndsu.edu/PHKG/patient1/medication1>
<http://schema.org/startDate> "2010-05-01"^^<http://www.w3.org/2001/XMLSchema#date> .
<https://cs.ndsu.edu/PHKG/patient1/medication1>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://cs.ndsu.edu/PHKG/Medication> .
<https://cs.ndsu.edu/PHKG/patient1>
<http://cs.ndsu.edu/PHKG/hasMedication>
<https://cs.ndsu.edu/PHKG/patient1/medication1> .

In this example, we have a patient identified by the URI "*https:// cs.ndsu.edu/PHKG /patient1*" who has a *name*, *birth date*, and *gender*. The patient also has a *condition* and a *medication*. Each of the entities in the graph (*patient*, *condition*, and *medication*) have a type specified by the RDF "*type*" predicate. The patient has a type of "*Person*", while the *condition* and *medication* have types specified by the ontology. The patient is linked to their condition and medication using predicates "*hasCondition*" and "*hasMedication*", respectively.

B. Authentication and Access Control

Blockchain is used to provide secure authentication and access control for personal health knowledge graphs. The whole system is powered by Ethereum, each user maintains a EOA, hence each user is uniquely identified by their unique Ethereum address. This identity can be used to verify the user's identity when accessing the personal health knowledge graph. Different users may have different roles, such as patient, doctor, insurance agency, etc. When a user logs in to the system, their digital identity is checked against the blockchain to ensure that they are authorized to access the system. In addition to user authentication, users' wearable devices can also be authenticated using blockchain. Each device is given a unique digital identity that is stored on the blockchain. When a device connects to the system, its digital identity is checked against the blockchain to ensure that it is authorized to access the system.

Once a user or sensing device is authenticated, access control policies can be enforced using smart contracts on the blockchain. These policies specify which parts of the personal

health knowledge graph each user or device is authorized to access. For example, a doctor may be authorized to access a patient's medical records, while a fitness tracker may only be authorized to access data related to exercise and physical activity.

For example, a doctor who is treating a patient with diabetes can query the personal health knowledge graph to retrieve the patient's blood sugar levels, medication history, and other relevant data. This information can help the doctor to make informed decisions about the patient's treatment plan and monitor their progress over time. To do that, the doctor initiates a request to access a patient's health data by creating a transaction on the blockchain network that invokes the authorization request function on the smart contract. The authorization request function takes in the Ethereum address of the doctor and the Ethereum address of the patient. The function checks if the doctor is authorized to access the patient's data by looking up the patient's permissions on the personal health data. If the doctor is authorized, the function creates an approve event and sends it to all the nodes on the blockchain network. The event contains a token which is hashed using the Ethereum address of the doctor and a nonce, the event also contains an expiration time for the token. The doctor then uses this token to send a signed message containing the token using the doctor's private Ethereum key to the resource system. The system can query the smart contract and get the details such as assigned token, role, and the doctor's Ethereum address. The signed message helps the system verifies the doctor's identity and to allow access to the system based on the role. The doctor can then query the personal health knowledge graph to retrieve the patient's health data that is relevant to their treatment.

Fig. 3 shows the sequence diagram of this communication process.

C. *Smart Contract*

In our blockchain-based PHKG network, smart contracts are used to automate the process of registration, authentication, and access control of individual users, health data producers, and data consumers. A smart contract is a self-executing code that runs on the blockchain network.

To register individual users, health data producers, and data consumers, we can create a smart contract that stores the necessary information, such as user identification, contact information, and access permissions. To authenticate people and institutes, we can use the Ethereum network's public-key infrastructure to verify user identities. The smart contract can include a function that verifies a user's identity by checking their public key against a stored list of authorized public keys. To control access to health data, we can use the smart contract to manage access permissions for users and data producers. The smart contract can include functions that allow authorized users to grant or revoke access permissions for specific data. For example, a function could be created to grant a data consumer access to a specific health data record.

In our system, we use the Ethereum blockchain platform to deploy smart contract. Our smart contract is written and tested with Solidity and Remix. It implemented the following functions:

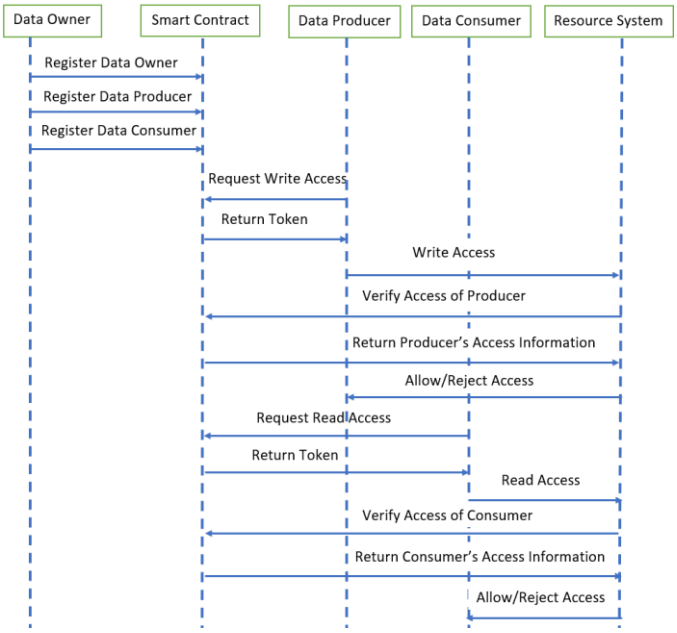

Fig. 3 A sequence diagram of the data access communication process

- registerDataOwner(): registers individual personal health data owners, i.e., person whose health data is being shared.
- registerDataProducer(): register institutes or devices, such as clinics, labs, doctors, and wearable devices, that produce health data.
- registerDataConsumer(): register the health data consumers such as doctors, insurance agents, data owner's family members, etc.
- getAccess_Producer():data producer execute the authorize function to obtain the authorization token from the smart contract.
- getAccess_Consumer(): data consumers execute the authorize function to obtain the authorization token from the smart contract.
- verifyAccess(): used by the resource system to get token, role, and access rules from the smart contract to verify a producer or consumer's identity.

Fig. 4 shows the smart contract written in Solidity language. The function getAccess_Consumer allows a data consumer to request authorization to access a patient's data. The function checks if the requesting data consumer is authorized by the patient to access their data by verifying if the requesting data consumer's address is present in the patientToDataUserMapping array for the patient's address. If the data consumer is authorized, a token is generated and associated with the data consumer's address and role, and emitted as an event using the tokenEvent function.

Smart contracts play a critical role in our system as they eliminate the need for intermediaries and provide a secure and transparent platform for data sharing. They automate the process of verifying and executing the contract terms, reducing the risk of fraud and ensuring that the terms are enforced as intended. Smart contracts also ensure that the personal health data is only accessible to authorized entities, which enhances data privacy and security. Overall, smart contracts provide a reliable and secure platform for the sharing of personal health data in a transparent and auditable manner.

```
//data consumers call this function to get token
function getAccess_Consumer(address _patientAddress)  public  {    infinite gas
    bool isRequestingDataConsumerAuthorized=false;

    for(uint i=0;i<patientToDataConsumerMapping[_patientAddress].length;i++){
        if(patientToDataConsumerMapping[_patientAddress][i]==msg.sender){
            isRequestingDataConsumerAuthorized=true;
        }
    }

    if(isRequestingDataConsumerAuthorized==true){
        bytes32 token=keccak256(abi.encodePacked(nonce,msg.sender));
        nonce++;

       dataConsumerDetails memory obj= dataConsumerAddressToConsumerDetailsMapping
       obj.token=token;
       dataConsumerAddressToConsumerDetailsMapping[msg.sender]=obj;

       //emit token here
       emit tokenEvent("Data Consumer",msg.sender,token,obj.role);

    }
    else{
        revert("Unauthorized Data Consumer");
    }
}
```

Fig. 4. A snippet of the proposed smart contract.

D. Distributed Graph Storage

The PHKG is broken down into smaller chunks of data, such

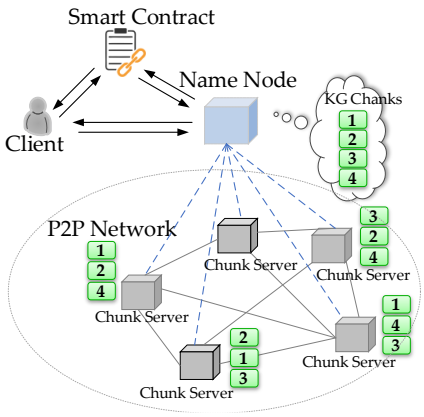

Fig. 5. P2P-based distributed PHKG storage

as triples, and each chunk can be assigned a unique identifier. As shown in Fig. 5, the chunks of data are then stored across multiple chunk servers, i.e., P2P nodes, to ensure fault tolerance and availability. The metadata for the PHKG is stored in the name node. This includes information such as the name of the graph, the size of the graph, and the location of the chunk servers. To facilitate faster retrieval of data, an DHT-based indexing is put in place to allow for fast lookup of data based on the unique identifiers assigned to each chunk. To ensure data durability, the data is replicated across multiple chunk servers, and backup mechanisms are used to recover from failures. Access control is implemented using the aforementioned smart contract implemented in the blockchain. P2P-based distributed storage can provide an efficient and scalable way to store a knowledge graph by distributing the data across multiple nodes in the network. This can improve fault tolerance, data availability, and data access speeds while ensuring data security and integrity.

IV. Evaluation

We deployed the proposed mechanism over an Ethereum blockchain network. In the first set of experiments, we see how blockchain and smart contracts protect the privacy of users' data and improve the system's security. Then we evaluate the performance of the decentralized access control, query, and storage system in terms of scalability, load balancing, and fault tolerance.

A. Access Control Verification

1) Smart contract use case testing

In this case, John is a patient at John Hopkins Hospital. He has opted to use a proposed blockchain-based personal health knowledge graph to securely and decentrally store and share all of his health data. As shown in Fig. 6, John has registered with the application using his Ethereum address (0xAb8…35cb2), and all his data generated during his hospital visits is safely stored and accessible through the application.

Fig. 6. John was successuflly registered as the data owner.

John has a new health provider Dr. Johnson with an Ethereum Address (0x4B2…C02db). John registers Dr. Johnson and provides him a doctor's role (Fig. 7).

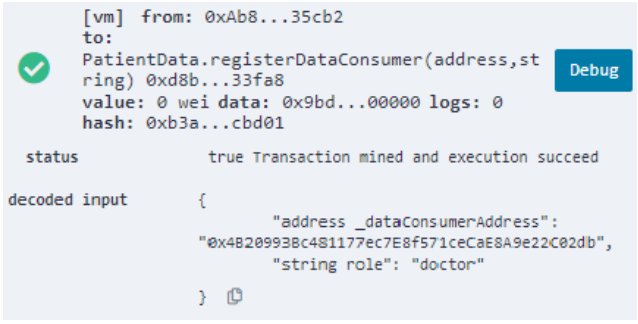

Fig. 7. Dr. Johnson was successuflly registered as the data consumer.

Dr. Johnson can access John's health data by interacting with the smart contract and receives a token, as shown in Fig. 8. The token obtained through the authorization process can be used by the doctor (authorized data consumer) to access the resource system where John's (patient) data is stored securely.

Fig. 8. Dr. Johnson was successuflly authrorized to accesss John's data.

An unauthorized party with Ethereum address (0x5c6…21678) tries to access John's data will be denied by the smart contract as shown in Fig. 9.

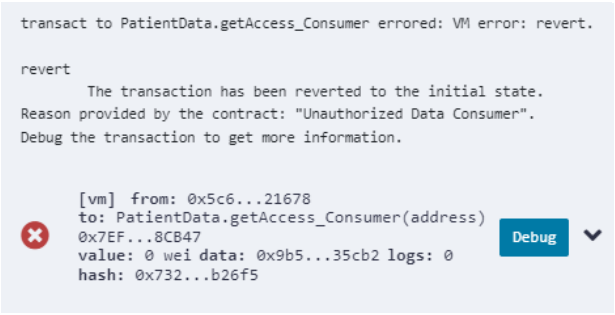

Fig. 9. Unauthorized party was denied with accesss of John's data.

2) Cost anlaysis

To understand the cost of smart contract operations, a prototype contract was created and deployed using Remix on the Ethereum test network. The cost of deploying the contract and executing the functions on the contract were analyzed. As shown in Table 1, transactions record the cost in gas in the Ethereum ecosystem. Gas is the fee required to execute a transaction on the Ethereum platform. The transaction cost varies depending on the complexity of the function being executed. The highest cost was associated with the contract creation transaction, while the function "verifyAccess" is a view function and does not cost anything if called externally. It is important to consider the cost of smart contract operations to ensure that the system is efficient and cost-effective to use for all users.

Table I. Transaction cost

Functions	**Gas Used**
Contract Creation	2445488
registerDataOwner	67837
registerDataConsumer	67837
getAccess_Consumer	62723
verifyAccess	0.0

B. *Security Analysis*

This section explains how our framework utilizes the blockchain to ensure the security of the entire system and prevent attacks.

- *Denial of Service attack:* All devices interacting with the smart contract in the Ethereum ecosystem require payments in the form of gas. This helps in restricting the flooding of requests to the smart contract and hence secures the platform from Denial-of-Service attack.
- *MITM and Replay Attack :* User sends token signed with private key to the resource system during interaction. Resource system deciphers token using user's public key to ensure authenticity. Public-private key ensures elimination of MITM and replay attacks in the system.
- *Integrity:* The smart contract on the Ethereum blockchain ensures data received from the contract is tamper-proof, providing evidence of integrity.
- *Authorization:* The smart contract on the Ethereum platform is also responsible for authorizing users in any off-chain communications to use any service. Users are given tokens that they use during off-chain communications. Additionally, solidity modifiers and logic have been added to restrict access and allow only valid devices to execute a particular function, guarding against the reentrancy attack.

C. *Network Performance*

We also deployed graph-based data over the P2P storage network. We measured the system's performance in terms of scalability, load balancing, and fault tolerance.

We compared the P2P distributed PHKG system's network communication overhead (bytes per node) with a centralized storage system. As illustrated in Fig. 10, as the number of users of the system increases, the system's overhead increases. This increase is dramatical for the centralized system. While the P2P nodes' overhead is much lower for P2P system with 50 nodes and 1000 nodes. For the P2P system with 1000 nodes the overhead is so small that its performance line almost touches the x-axis. These experiments demonstrate the scalability of the P2P system.

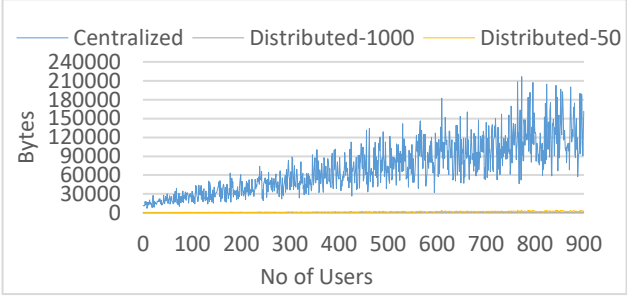

Fig. 10. Number of users (patients) vs. avg message overhead per server (node).

Fig. 11 shows the distribution of communication load of a P2P network with 500 nodes. We randomly generated data access requests. For simplicity, we only generate "read" requests. As shown in the figure, communication overhead is basically balanced among the P2P nodes mainly ranging from

10000 to 30000 bytes. These experiments illustrate the load balancing properties of the system.

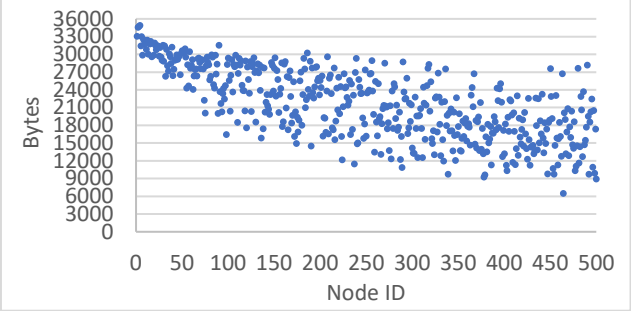

Fig. 11. Nodes' communication overhead distribution

As presented in the previous section, to ensure data availability, in the P2P-based distributed PHKG storage, data is replicated across multiple chunk servers. We tested the data availability of the system. In this experiment, each data has three replicas, i.e., each data chunk is replicated in three P2P nodes. The network size is 500. Fig. 12 shows the success rate of the access requests with respect to the node drop rate. The data availability keeps being 100% when the node's drop rate is below 30%. For the network of 500 nodes, even 150 nodes die or leave the system, the network still can provide 100% data availability. These experiments show the fault tolerance of the system.

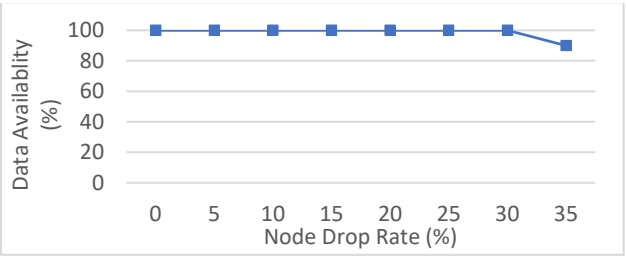

Fig. 12. Data availability of the PHKG storage vs the network's node drop rate

V. Conclusions

The amount of health-related data generated from various sources continues to increase, providing valuable insights into an individual's health. However, managing and integrating this data poses significant challenges, including data interoperability, privacy, and security. To address these challenges, we propose a personal health knowledge graph that organizes PHD in a graph format, enabling the integration and analysis of various types of data for a comprehensive view of an individual's health. Furthermore, we suggest a blockchain architecture to provide a secure and tamper-proof way of storing and sharing PHKG, allowing patients to have full control over their data and to share it with trusted entities as needed. The use of smart contracts also enables the automation of access control and data sharing policies, providing enhanced privacy and security. The proposed system has been evaluated using use cases and simulation, and it has been shown to be secure, scalable, and enables effective data sharing and integration.

References

[1] T. Justinia, "Blockchain Technologies: Opportunities for Solving Real-World Problems in Healthcare and Biomedical Sciences," *Acta Informatica Medica*, vol. 27, no. 4, p. 284, 2019

[2] A. Haleem, M. Javaid, R. P. Singh, R. Suman, and S. Rab, "Blockchain technology applications in healthcare: An overview," International Journal of Intelligent Networks, vol. 2, pp. 130–139, Jan. 2021.

[3] A. Haleem, M. Javaid, R. P. Singh, R. Suman, and S. Rab, "Blockchain technology applications in healthcare: An overview," *International Journal of Intelligent Networks*, vol. 2, pp. 130–139, Jan. 2021.

[4] H. A. Lee *et al.*, "An Architecture and Management Platform for Blockchain-Based Personal Health Record Exchange: Development and Usability Study," *J Med Internet Res 2020;22(6):e16748 https://www.jmir.org/2020/6/e16748*, vol. 22, no. 6, p. e16748, Jun. 2020,

[5] A. Cernian, B. Tiganoaia, I. S. Sacala, A. Pavel, and A. Iftemi, "PatientDataChain: A Blockchain-Based Approach to Integrate Personal Health Records," *Sensors 2020, Vol. 20, Page 6538*, vol. 20, no. 22, p. 6538, Nov. 2020, doi: 10.3390/S20226538.

[6] M. Wazid, A. K. Das, and Y. Park, "Blockchain-enabled secure communication mechanism for IoT-driven personal health records," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 4, p. e4421, Apr. 2022, doi: 10.1002/ETT.4421.

[7] M. Chen *et al.*, "Blockchain-Enabled healthcare system for detection of diabetes," *Journal of Information Security and Applications*, vol. 58, p. 102771, May 2021, doi: 10.1016/J.JISA.2021.102771.

[8] K. M. Malik, M. Krishnamurthy, M. Alobaidi, M. Hussain, F. Alam, and G. Malik, "Automated domain-specific healthcare knowledge graph curation framework: Subarachnoid hemorrhage as phenotype," *Expert Syst Appl*, vol. 145, p. 113120, May 2020,

[9] W. Choi and H. Lee, "Identifying disease-gene associations using a convolutional neural network-based model by embedding a biological knowledge graph with entity descriptions," *PLoS One*, vol. 16, no. 10, p. e0258626, Oct. 2021, doi: 10.1371/JOURNAL.PONE.0258626.

[10] Y. Zhang *et al.*, "HKGB: An Inclusive, Extensible, Intelligent, Semi-auto-constructed Knowledge Graph Framework for Healthcare with Clinicians' Expertise Incorporated," *Inf Process Manag*, vol. 57, no. 6, p. 102324, Nov. 2020, doi: 10.1016/J.IPM.2020.102324.

[11] L. Shi, S. Li, X. Yang, J. Qi, G. Pan, and B. Zhou, "Semantic Health Knowledge Graph: Semantic Integration of Heterogeneous Medical Knowledge and Services," *Biomed Res Int*, vol. 2017, 2017, doi: 10.1155/2017/2858423.

[12] X. Tao *et al.*, "Mining health knowledge graph for health risk prediction," *World Wide Web*, vol. 23, no. 4, pp. 2341–2362, Jul. 2020, doi: 10.1007/S11280-020-00810-1/TABLES/8.

[13] L. Li *et al.*, "Real-world data medical knowledge graph: construction and applications," *Artif Intell Med*, vol. 103, p. 101817, Mar. 2020, doi: 10.1016/J.ARTMED.2020.101817.

[14] V. K. C. Yan *et al.*, "Drug Repurposing for the Treatment of COVID-19: A Knowledge Graph Approach," *Adv Ther (Weinh)*, vol. 4, no. 7, p. 2100055, Jul. 2021, doi: 10.1002/ADTP.202100055.

[15] S. K. Mohamed, V. Nováček, and A. Nounu, "Discovering protein drug targets using knowledge graph embeddings," *Bioinformatics*, vol. 36, no. 2, pp. 603–610, Jan. 2020, doi: 10.1093/BIOINFORMATICS/BTZ600.

[16] M. Samwald *et al.*, "Linked open drug data for pharmaceutical research and development," *J Cheminform*, vol. 3, no. 1, May 2011, doi: 10.1186/1758-2946-3-19.

[17] V. Pandey, J. Li, and S. Alian, "Evaluation and Evolution of NAOnto - An Ontology for Personalized Diabetes Management for Native Americans," *2021 7th International Conference on Computer and Communications, ICCC 2021*, pp. 1635–1641, 2021,

[18] S. Alian, J. Li, and V. Pandey, "A Personalized Recommendation System to Support Diabetes Self-Management for American Indians," *IEEE Access*, vol. 6, pp. 73041–73051, 2018,

[19] " FHIR v4.3.0." https://hl7.org/fhir/overview.html (accessed Mar. 09, 2023).

[20] V. Kilintzis, A. Kosvyra, N. Beredimas, P. Natsiavas, N. Maglaveras, and I. Chouvarda, "A sustainable HL7 FHIR based ontology for PHR data⁎," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 5700–5703, Jul. 2019, doi: 10.1109/EMBC.2019.8856415.