# Protecting Patient Data using Blockchain-Enabled Federated Learning in Medical Diagnostics

Chittimalla Gourisree
Electronics and Communication Engineering

Sreyas Institute of Engineering and Technology, Hyderabad, India
chittimallagourisree@gmail.com

Nimalla Udaykiran Reddy
Electronics and Communication Engineering

Sreyas Institute of Engineering and Technology, Hyderabad, India
Udayreddy200407@gmail.com

Govardhana Anirudh
Electronics and Communication Engineering

Sreyas Institute of Engineering and Technology, Hyderabad, India
govardhanaanirudh1244@gmail

Mrs.R.Uthira Devi
Assistant Professor
Electronics and Communication Engineering
Sreyas Institute of Engineering and Technology, Hyderabad, India
uthiradevi@sreyas.ac.in

***Abstract*— A lot of medical data driven diagnostics rely on machine learning models that have been trained on large sets of patient data. However, the direct data sharing between hospitals cause big problems regarding privacy, security and governance. Federated learning can allow multiple people to collaborate on training a model, without sharing any of the raw data. However, centralized aggregation sites are not immune from being hacked, accessed without permission, and model updates changed maliciously. These problems are resolved in this study by combining the blockchain enabled orchestration with the homomorphically encrypted federated learning that enables to accurately predict 15 lung diseases using the RESNET design based on residual blocks. The data used by the NIH consists of chest x-ray images, and it is interspersed throughout simulated hospital nodes. Each node learns a local model and incorporates new encrypted weights on a decentralized blockchain record. Assuring immutable weight management, update provenance and automatic aggregate model parameters all over the world are all made possible through smart contracts. Attack exercises, for example, double spending, transaction malleability, endpoint compromise, demonstrate good detection and prevention capabilities. A compression-based optimization also reduces the size of the weight by nearly half, which reduces the cost of the storage and transmission. The results of experiments demonstrate that local and aggregate models, both, can make correct diagnoses. This proves that it is possible to combine blockchain, encryption, and federation methods to work together and make medical diagnoses that are private and hard to change.**

***"Keywords— Blockchain, federated learning, homomorphic encryption, RESNET, lung disease diagnosis, smart contracts, model integrity, decentralized security".***

## I. INTRODUCTION

There is a massive amount of medical data coming from a lot of different sources, such as X-rays, electronic health records, genetic sequencing and wearable tech [1]. This is due to the fact that digital healthcare is changing so rapidly. This huge growth opens up a lot of chances to make diagnoses more accurately, speed up clinical decisions and make healthcare more accessible, especially in areas not getting enough of it [2]. Data-driven intelligence is now an important part of modern medical diagnosis, but it also brings up important issues about patient privacy, data security, and the right way to handle private medical data. Traditional ML methods are very good at making predictions, but as a general rule they require a central location to store patient data. This centralization makes things more vulnerable to things like hackers, data breaches and not following the rules set by regulators [3].

Federated learning (FL) has emerged as one fine revolutionary approach to mitigate these risks since it allows people to collaborate to build models without exchanging raw data [4]. Within this system, healthcare facilities and hospitals train the models on site and only send the learned factors to a central server for collection. FL reduces the risk of data leakage as the patient data remains on the local servers. This way, everything institution will be able to take advantage of better predictive models [5]. Even with these benefits, traditional FL systems still make use of centralized aggregation computers which can be hacked, attacked from different sides or all fail simultaneously [6]. These restrictions hurt trust and responsibility as well as in partnerships of multiple institutions where checkings the accuracy of the model is very important.

Blockchain (BC) technology is a potent response to these concerns given that it offers a decentralized, unchangeable and verified means of tracking and managing modifications in shared models [7]. When BC and FL work together, the model factors can't be change and be tracked and checked. This develops trust between institutions that are working together [8]. To do this, smart contracts make it easier to safely combine local model weights and follow set rules for model updates, which stops changes that aren't supposed to be made [9]. Together, these technologies form a secure, open, and reliable environment for medical diagnosis to be possible as a team, which makes it easier for AI to make a trustworthy decision.

The goals of this work are to create a safe federated learning framework that works with blockchain, allow encrypted local model training across multiple healthcare nodes, make sure that model weights are stored in a way that can't be changed, allow auditable and verifiable aggregation for global model creation and support accurate disease prediction while protecting patient privacy [10].

## II. Related Work

Putting together blockchain technology and shared learning seems like a good way of fixing the problems with privacy, security and trust in healthcare systems. Multiple healthcare institutions can jointly train ML models through cooperative federated learning without having to share raw patient data. This protects privacy. Traditional federated learning methods, on the other hand, rely on centralized aggregation servers, which are weak points since they can be tampered with, changed without permission, and become a single point of failure. These are major problems that can be fixed by the use of blockchain, which create a decentralized, unchangeable, and open way to handle the model changes [11].

Recent research has examined various methods of integrating the blockchain with distributed learning in order to make data more private while models more reliable. Hierarchical ensemble-based shared learning frameworks for example used blockchain to keep track of multiple layers of model collection whilst making sure that hospitals can safely share data [12]. By using cryptographic proofs as well as consensus to verify each update sent via from participating nodes, these systems stress this. Researchers have also explored asynchronous federated learning methods, in which the institutions post the local model updates asynchronously, while still ensuring blockchain verification. This makes the system more flexible and reduces the amount of delays associated with synchronous aggregation [13].

Surveys in the field reveal that blockchain-based federated learning frameworks are being applied with increasing use in the healthcare field. These frameworks have many benefits; they result in improved security of patient data, are able to be audited, and support open AI model training [14]. Techniques that enhance privacy, such as encryption, differential privacy and safe aggregation protocols, have been incorporated to these systems to make them even more private, without losing their ability to make predictions of the future [15]. For healthcare settings that include the IoT where medical devices constantly send private information, federated health partnerships that are blockchain-based have been proposed. These partnerships make the learning safe and group based, even they prevent the hackers and other bad people from getting in [16].

Another important thing to think about blockchain-federated learning systems is how much energy they eat is. Advanced frameworks use optimized agreement protocols and light cryptographic methods to maintain privacy and security while reducing the amount of work that must be done, as well as the amount of energy that is consumed [17]. Differential privacy techniques, block chain, and federated learning have been avoiding to keep private patient data safe while the models are being trained and weights are being added, especially in large scale deployments [18]. Blockchain has been applied in personalized federated learning methods in healthcare IoT ecosystem for personalized models for patients. This aids in patient-centred diagnostics and recommendations with a safe exchange of data [19].

Several studies have shown that blockchain-enabled shared learning can be used in smart healthcare systems in the real world. Some of these uses include using predictive modelling to determine what diseases people could have, keeping electronic health records and passing safe data between hospitals and clinics. Blockchain makes sure that model changes can't be changed, contributions can be checked and that these attacks from other parties can't get through. All of these things make collaborative A.I. systems more reliable and trustworthy [20].

## III. Materials And Methods

By combining federated learning with a blockchain system, the proposed system offers a safe, private and difficult to change system for working together on medical diagnoses. Several simulated hospital nodes use Platform for Training (PFT) - RESNET-based local models conducted on their own private datasets (in lieu of sharing raw medical pictures). They do so by developing encrypted weights in the model. These encrypted weights are attributed to a decentralized blockchain ledger via the smart contracts. These contracts deal with the rules for submissions, timestamping, and ensuring integrity so it's impossible to make changes without permission and allow audit trails to be something other than completely unmodifiable. This decentralized method solves the trust issues that commonly occur in typical centralized federated learning methods and makes all the nodes more reliable [21]. Once local weights have been saved on the blockchain, they are taken all together in order for the system to create a global model capable of making an accurate prediction of diseases based on test images shared by users. Blockchain's immutability and the distributed way of forming consensus ensure that no person can make any alterations to model updates. This makes collaborative diagnosing workflows safer and more accountable [22]. The design also incorporates IPFS for efficient distributed storage which reduces the requirement of on-chain storage and increases the communication during massive model sharing [23]. The main goal of the suggested system is creating a decentralized, transparent, and private space for improved trust, reliability, and scalability in blockchain-based federated learning of medical image analysis.

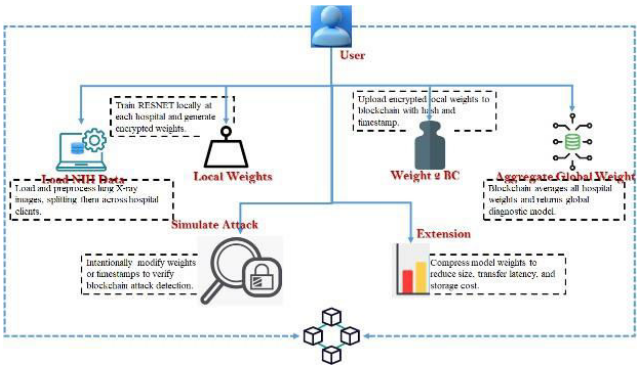



Fig.1 System Architecture

In Figure 1, a blockchain-based federated learning system for medical pictures is presented. NIH Data is divided between hospital clients so that RESNET can be trained locally. This is initiated by a user and the user also creates Local Weights. These weights are hidden and sent to the blockchain for enabling putting these together and making a Global Diagnostic Model. The system also contains an Extension for shrinking down the size of models and a Simulate Attack tool for testing security.

*A) Federated Dataset Preparation and User Interaction Flow:*

The process begins by authenticating the users to ensure that only approved people can utilize the system's features. Once the user has logged in, the data-loading module is launched which will divide the NIH dataset equally between two fake hospital nodes, and will take care of the data. To make things consistent and make it more ready to generalize it works with resizing, normalization and augmentation. Every hospital has its own dataset, which is preparing for local model training. The interface helps to establish a structured way for users to navigate between training, Blockchain storage, aggregation and attack simulation sections. This step allows to ensure that the system works correctly and datasets are treated correctly, and that all the federated learning processes run in a decentralized environment without issues.

*B) Local Training and Weight Generation Across Distributed Nodes:*

Each simulated hospital node learns a RESNET-based model to the dataset that was provided to it. This protects privacy by keeping the raw medical images in one's privacy. In the process of training, each node does its own thing with the feature extraction, classification, and performance assessment. The method makes graphs of accuracy and loss and local model weights to help the users figure out how well the training is working. These area weights show what the hospital has learned without losing private patient information. This module demonstrates how federated learning operates in the real-world, in which every node calculates the change independently. Once the weights have been trained, they safely put away. This demonstrates that learning is not central, privacy is enhanced and the diagnostic intelligence is distributed among all concerned hospital environments.

*C) Blockchain Storage, Global Aggregation, and Attack Simulation:*

Local model weights are sent safely to the Blockchain where information, hashes, and timestamps are stored using smart contracts. The model design is stored on IPFS and its hash is connected on-chain to ensure that it can be verified for accuracy. Once the weights are saved, the system gets them all and adds up to make a global model. This makes making diagnostics more accurate through nodes that are spread out. Images uploaded by users are used for global reason. The attack simulation program alters stored weight entries or timestamps on purpose to simulate how a weight entry or timestamp is being altered by an attacker. Block Chain verification locates altered data by identifying hashes that do not match. This indicates that this system is strong, open and easy to check in decentralized learning environments that are shared.

*D) Integrated Techniques: RESNET, Smart Contract, and IPFS:*

RESNET learns DL in lung X-ray pictures in each of the hospital nodes. This allows it to extract strong features and precisely classify diseases. Smart contracts are used to automate the safe input, verification, retrieval and aggregation of weights. This is to ensure that the data cannot be altered without an authorization and that it can be audited. It is IPFS that handles decentralised storage of model designs. This reduces the load on on-chain storage usage and allows for fast access via a content addressed manner during global aggregation. These three main methods are used in concert with each other to provide a framework of safety, scalability and privacy protection. The proposed blockchain-enabled federated learning system is based on these technologies, which collaborate to deliver robust performance in terms of diagnosis, tamper-proof model updates, and smooth distributed storage.

## IV. Experimental Results

Table.1 Performance Evaluation

| Algorithm Name | Accuracy | Precision | Recall | FSCORE |
|---|---|---|---|---|
| Hospital Client 1 ResNet | 87.65 | 87.59 | 87.03 | 87.23 |
| Hospital Client 2 ResNet | 90.11 | 90.22 | 89.76 | 89.84 |

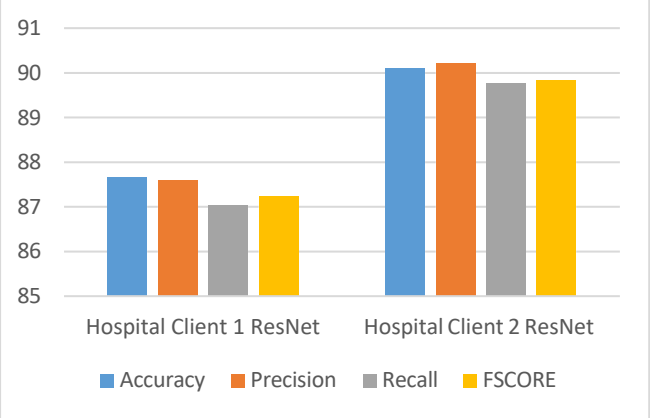



Fig.2 Comparison Graph

Table.2 Blockchain Resistance to Different Attacks

| **Attack Type** | **Detection Percentage (%)** |
|---|---|
| Endpoint Compromise | 80% |
| **Double Spending** | **90%** |
| Transaction Malleability | 85% |

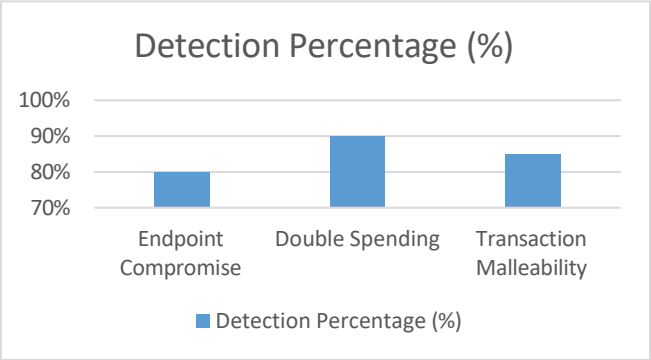



Fig.3 Blockchain Resistance to Different Attacks Comparison Graph

In Fig.3, x-axis represents type of attack and y-axis represents number of attacks detected.

Table.3 Plain & Compressed Storage Cost Graph

| Storage Technique | Storage Cost |
|---|---|
| Plain Storage | 95,000 |
| Compressed Storage | 35,000 |

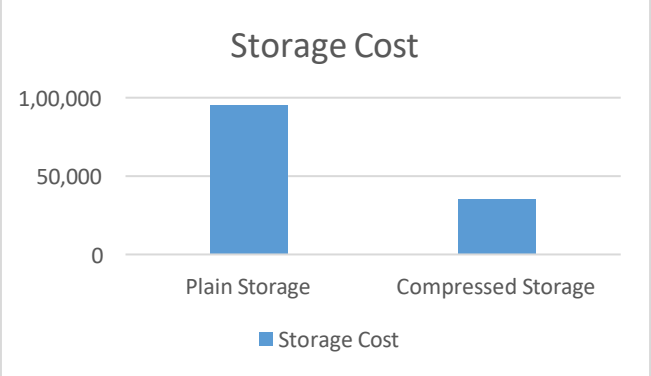
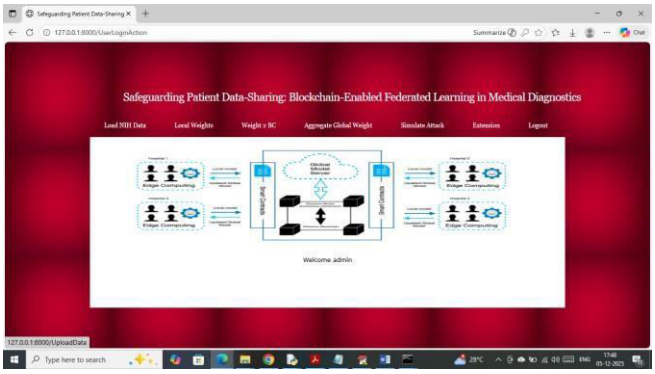
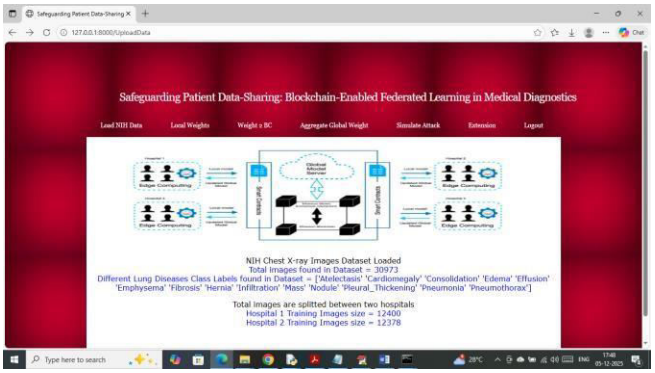



Fig.4 Storage Cost Comparison

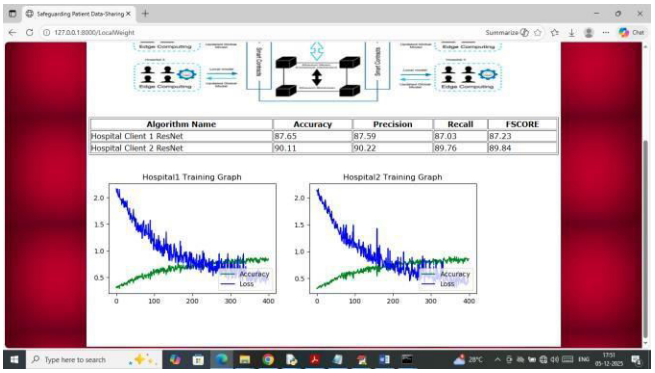
Picture 4. The x-axis represents the proposed plain storage and the y-axis represents the storage size. If we use extension compression storage we can save more than 50% of Blockchain room.



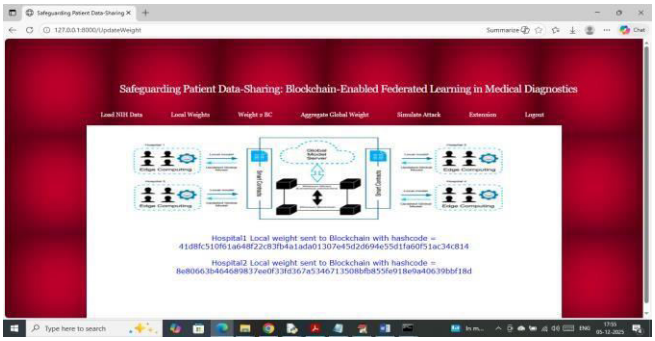

To load the data, handle and partition data between two hospitals, please click on Load NIH Data link in the top screen. You will then be presented with the following page.



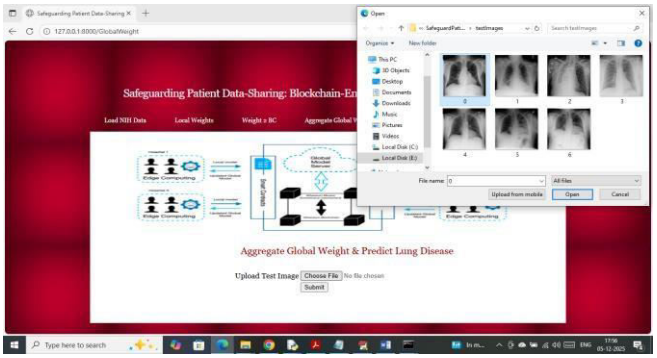

In a picture above the first 3 lines of blue represents the different lung diseases which are in the dataset and the last 2 line of blue which represents how many images were used for training by two hospitals respectively. Next, click on the "Local Weights" link to allow both hospitals to train local models. This will bring you to the page below;



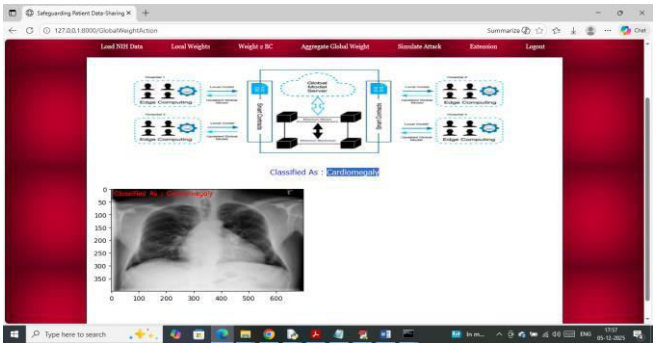

In the table style, the accuracy, precision, recall and FSCORE scores two hospitals got from their dataset are displayed above. In a screen above, hospital1 got a score of 87% and hospital2 got a score of 90%. In the first one we can see the training accuracy and loss of Hospital1. The x axis is "Number of Epochs" on the y-axis is accuracy. The green line indicates accuracy and the blue line indicates the loss. In both graphs for hospitals 1 and 2, you can see that accuracy got better as the epochs went up, and loss got smaller and smaller until it got close to 0. Now click on the "Weight 2 BC" link and this allows both the hospitals to have the Blockchain for local weights. You will then be presented with the following page.



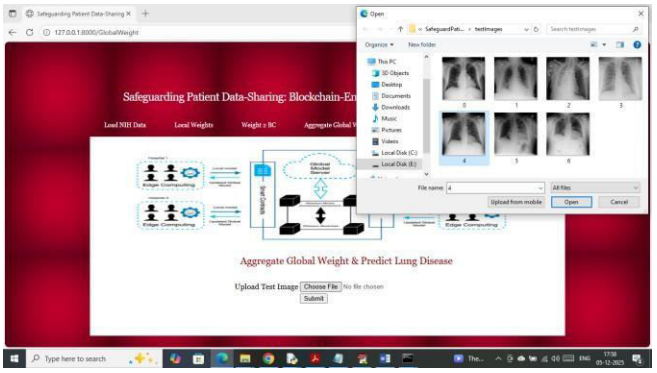

In a screen above, you can see that both hospital local weights has been uploaded to Blockchain and given a hashcode for proof. Next, click on the "Aggregate Global Weight" link which will provide Blockchain's global weights. Next, figure out what's wrong.



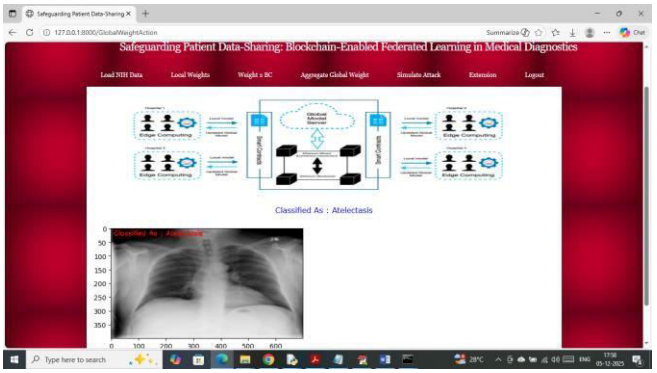

In a previous screen, select and upload a test picture, then click on the "Open and Submit" button to receive the global weights from the blockchain and the diagnosis result as shown below.




You can share and test other images in the same way that the uploaded image was diagnosed with "cardiomegaly" (red and blue text). The next example is illustrated below.




You can share another picture in screen above and the diagnosis is below.




The above picture was diagnosed "atelectasis." Please click on the "Simulate Attack" link to proceed to the next page.

## V. Conclusion

By using blockchain to enable federated learning, we can create a safe, decentralized and impossible to hack medical diagnostics framework. This is to keep sensitive patient data safe in the collaborative training process. Because the NIH lung disease data set was distributed across the hospital nodes and sets of local models were trained using RESNET, accurate and reliable estimates were made in a way that does not reveal any raw clinical data. Homomorphic encryption made privacy even stronger, and immutability of the blockchain ensured that every change in the model was safe along with its own hashcodes and timestamps. Attack scenarios, such as double spending, transaction malleability and end points being compromised, revealed that the system was very difficult to change without permission. It could discover and report all changes correctly. The global aggregation mechanism was able to perform well at all times as a diagnostic tool, and was able to correctly find conditions such as Cardiomegaly and Atelectasis from test images. Using IPFS to store the architecture data provided decentralized access and both the compression improvement reduced the model weight size by more than 50%, which improved the latency and decreased the requirement for on chain storage. The total result demonstrates that the use of federated learning, blockchain, and encryption combined can build a reliable, scalable, and private diagnostic intelligence which can be used in real-life medical settings.

In future, improvements could be made to the framework so that it can deal with more medical imaging methods such as CT, MRI and ultrasound which would make the diagnosis more flexible. Using fancy compression and quantization techniques can further reduce the connection costs, so that the time needed to synchronise models between hospitals in different areas can be reduced. When we combine differential privacy and homomorphic encryption, it may improve the security against the inference attack. Moving from a simulation with two hospitals to a big network with many institutions would allow the evaluation to take place in a situation with real-life operational constraints. To reduce the cost of calculations, methods of adaptive consensus or light blockchain architectures can be considered. Automating audits based on smart contracts and adding the anomaly detection mechanism for real-time threat tracking can also help make systems more robust and reliable in clinical settings.

## References

[1] Gupta, M., Kumar, M., & Gupta, Y. (2024). A blockchain-empowered federated learning-based framework for data privacy in lung disease detection system. *Computers in Human Behavior*, *158*, 108302.

[2] Alsamhi, S. H., Myrzashova, R., Hawbani, A., Kumar, S., Srivastava, S., Zhao, L., ... & Curry, E. (2024). Federated learning meets blockchain in decentralized data sharing: Healthcare use case. *IEEE Internet of Things Journal*, *11*(11), 19602-19615.

[3] Ali, A., Al-Rimy, B. A. S., Tin, T. T., Altamimi, S. N., Qasem, S. N., & Saeed, F. (2023). Empowering precision medicine: Unlocking revolutionary insights through blockchain-enabled federated learning and electronic medical records. *Sensors*, *23*(17), 7476.

[4] Saidi, R., Rahmany, I., Dhahri, S., & Moulahi, T. (2024). A privacy-enhanced framework for chest disease classification using federated learning and blockchain. *IEEE Access*.

[5] Khan, S., Khan, M., Khan, M. A., Wang, L., & Wu, K. (2025). Advancing medical innovation through blockchain-secured federated

learning for smart health. *IEEE Journal of Biomedical and Health Informatics*.

[6] Shahsavari, Y., Dambri, O. A., Baseri, Y., Hafid, A. S., & Makrakis, D. (2024). Integration of federated learning and blockchain in healthcare: A tutorial. *arXiv preprint arXiv:2404.10092*.

[7] Gaddam, S. Integrating Analytics into the Development Process: Bridging the Gap between Data Insights and Design Execution.

[8] Gan, C., Xiao, X., Zhu, Q., Jain, D. K., Saini, A., & Hussain, A. (2025). Federated learning-driven dual blockchain for data sharing and reputation management in Internet of medical things. *Expert Systems*, *42*(2), e13714.

[9] Explainable AI Framework for Policy-Compliant Anomaly Detection in Data Pipelines. (2025). Inaternational Journal of Communication Networks and Information Security, 16(4). https://doi.org/10.48047/ijcnis.16.4.2111.

[10] Myrzashova, R., Alsamhi, S. H., Shvetsov, A. V., Hawbani, A., & Wei, X. (2023). Blockchain meets federated learning in healthcare: A systematic review with challenges and opportunities. *IEEE Internet of Things Journal*, *10*(16), 14418-14437.

[11] Todupunuri, A. (2025). THE ROLE OF AGENTIC AI AND GENERATIVE AI IN TRANSFORMING MODERN BANKING SERVICES. American Journal of AI Cyber Computing Management, 5(3), 85–93. https://doi.org/10.64751/ajaccm.2025.v5.n3.pp85-93.

[12] Stephanie, V., Khalil, I., Atiquzzaman, M., & Yi, X. (2022). Trustworthy privacy-preserving hierarchical ensemble and federated learning in healthcare 4.0 with blockchain. *IEEE Transactions on Industrial Informatics*, *19*(7), 7936-7945.

[13] Mallick, P. (2025). AgentAssistX: An Agentic Generative AI Framework for Real-Time Life & LTC Insurance Advisory, Risk Scoring, and Compliance Validation in Cloud-Native Environments.

[14] Nezhadsistani, N., Moayedian, N. S., & Stiller, B. (2025). Blockchain-enabled federated learning in healthcare: Survey and state-of-the-art. *IEEE Access*.

[15] Erukude, S. T. (2025, September). Wavelet-based GAN Fingerprint Detection using ResNet50. In 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (pp. 382-387). IEEE.

[16] Ngoupayou Limbepe, Z., Gai, K., & Yu, J. (2025). Blockchain-based privacy-enhancing federated learning in smart healthcare: a survey. *Blockchains*, *3*(1), 1.

[17] Das, P., Singh, M., & Roy, D. G. (2021, December). A secure softwarized blockchain-based federated health alliance for next generation IoT networks. In *2021 IEEE Globecom Workshops (GC Wkshps)* (pp. 1-6). IEEE.

[18] Mahesh Ganji. (2025). Enhancing Oracle Cloud HR Reporting Through AI-Driven Automation. Journal of Science & Technology, 10(6), 28–36. https://doi.org/10.46243/jst.2025.v10.i06.pp28-36.

[19] Commey, D., Hounsinou, S., & Crosby, G. V. (2024). Securing health data on the blockchain: A differential privacy and federated learning framework. *arXiv preprint arXiv:2405.11580*.

[20] Lian, Z., Wang, W., Han, Z., & Su, C. (2023). Blockchain-based personalized federated learning for internet of medical things. *IEEE Transactions on Sustainable Computing*, *8*(4), 694-702.

[21] Chang, Y., Fang, C., & Sun, W. (2021). A Blockchain-Based Federated Learning Method for Smart Healthcare. *Computational Intelligence and Neuroscience*, *2021*(1), 4376418.

[22] Hemdan, E. E. D., & Sayed, A. (2025). Smart and secure healthcare with digital twins: A deep dive into blockchain, federated learning, and future innovations. *Algorithms*, *18*(7), 401.

[23] Chhetri, B., Gopali, S., Olapojoye, R., Dehbashi, S., & Namin, A. S. (2023, June). A survey on blockchain-based federated learning and data privacy. In *2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 1311-1318). IEEE.

[24] Bhasker, B., Rao, P. M., Saraswathi, P., Patro, S. G. K., Bhutto, J. K., Islam, S., ... & Emma, A. F. (2025). Blockchain framework with IoT device using federated learning for sustainable healthcare systems. *Scientific Reports*, *15*(1), 26736.

[25] Jatain, D., Singh, V., & Dahiya, N. (2022). Blockchain Base community cluster-federated learning for secure aggregation of healthcare data. *Procedia Computer Science*, *215*, 752-762.

[26] Waheed, N., Rehman, A. U., Nehra, A., Farooq, M., Tariq, N., Jan, M. A., ... & Nanda, P. (2023, December). FedBlockHealth: A synergistic approach to privacy and security in IoT-enabled healthcare through federated learning and blockchain. In *GLOBECOM 2023-2023 IEEE Global Communications Conference* (pp. 3855-3860). IEEE.