

A Blockchain-Based Framework for Enhancing Data Integrity and Privacy Protection in Cloud Storage Networks

PRATYUSH SINGH¹

¹C-9, 9796, VASANT KUNJ, NEW DELHI, 110070

Abstract

Cloud storage networks have become the default infrastructure for retaining sensitive digital records, yet their reliance on centralized providers exposes data to silent tampering, unauthorized disclosure, and single points of failure. This paper presents a blockchain-based framework that jointly enhances data integrity and privacy protection in cloud storage networks, using healthcare records as the motivating application. The framework is organized into three layers: an access layer through which patients and providers interact with a client application, a control layer in which a permissioned consortium blockchain of five institutions governs four dedicated chaincodes (EHRChain, DiagnosisChain, PolicyChain, and LogChain), and a storage layer built on the InterPlanetary File System (IPFS). Encrypted records are stored off-chain, while only their SHA-256-derived content identifiers and ownership proofs, represented as non-fungible tokens (NFTs), are anchored on-chain. To support collaborative analytics without exposing raw data, the framework integrates blockchain-validated federated learning, in which clients preprocess data with min-max normalization, select features with a Levy Grasshopper Optimization Algorithm (LGOA), train locally, and submit hashed model updates that are validated before federated averaging with explainable AI (XAI). An Entropy Deep Belief Network (EDBN) performs attack detection and disease prediction. Experiments on the Heart Disease and Breast Cancer Wisconsin datasets show 100% detection of injected tampering, 99.8% rejection of unauthorized access, disease prediction accuracy of 93.7% and 98.2%, and an overall attack detection recall of 97.9%, at a moderate latency overhead relative to a centralized store.

Keywords: Blockchain, Data Integrity, Privacy Protection, Cloud Storage, IPFS, Consortium Blockchain, Federated Learning, Deep Belief Network, Smart Contract.

1. Introduction

Cloud storage has become the backbone of modern digital services. Hospitals, banks, research institutions, and public agencies increasingly outsource the retention of sensitive records to cloud storage networks because of their elasticity, low upfront cost, and global reach. This convenience, however, comes with a structural weakness: the data owner surrenders physical control of the data to a provider whose internal behavior cannot be directly observed. A provider, an insider, or an external adversary who gains access to the storage back end can modify, delete, replay, or disclose records without leaving evidence that the owner can independently verify. Reported incidents of silent data corruption, misconfigured storage buckets, and credential theft show that integrity and privacy cannot be treated as guarantees that the cloud will simply honor.

Blockchain technology offers a different trust model. A distributed ledger replicated across mutually distrusting participants provides an append-only, tamper-evident record whose state is agreed upon by consensus rather than dictated by one operator. Smart contracts, known as chaincode in Hyperledger Fabric, allow access policies and audit rules to be enforced automatically and transparently [3]. Yet blockchains are poorly suited to storing bulk data. Writing large files directly on-chain inflates ledger size, increases latency, and exposes content to every node. A practical framework must therefore separate where data is kept from where trust is established: bulky encrypted data is stored in a decentralized off-chain network such as the InterPlanetary File System (IPFS) [4], while compact cryptographic evidence about that data is

anchored on the blockchain.

A further requirement arises in data-intensive domains such as healthcare. Institutions do not only wish to store records safely; they also wish to learn from them collectively, for example to build disease prediction models, without moving raw records between sites. Federated learning (FL) addresses this need by training models locally and sharing only parameters [5]. However, a federated process introduces its own integrity risks, because a malicious or compromised client can submit poisoned updates, and a central aggregator can be impersonated or can tamper with the global model. Anchoring model-update hashes on a blockchain and validating them before aggregation closes this gap, but it requires a careful co-design of the storage layer, the ledger, and the learning pipeline.

This paper proposes a blockchain-based framework that enhances data integrity and privacy protection in cloud storage networks while supporting privacy-preserving collaborative analytics. The framework is organized into three layers, as illustrated in Fig. 1. The access layer lets patients and healthcare providers register, generate records, and manage privacy preferences through a client application. The control layer is a permissioned consortium blockchain operated by five institutions (H1 to H5), whose validators execute four dedicated chaincodes. The storage layer is an IPFS off-chain network that holds only encrypted data. Proof of ownership is represented by a non-fungible token (NFT) that binds a data owner to a SHA-256-derived content identifier, as shown in Fig. 2. Finally, the analytics pipeline in Fig. 3 combines min-max normalization, Levy Grasshopper Optimization Algorithm (LGOA) feature selection, blockchain-validated federated averaging with explainable AI (XAI), and an Entropy Deep Belief Network (EDBN) for attack detection and disease prediction.

The principal contributions of this work are summarized as follows:

- A three-layer architecture that separates access, control, and storage responsibilities and keeps raw data off-chain while anchoring only hashes, ownership proofs, policies, and audit entries on a consortium blockchain.
- A four-chaincode design (EHRChain, DiagnosisChain, PolicyChain, and LogChain) that enforces patient-defined privacy preferences and records every access in an immutable audit trail.
- An NFT-based proof-of-ownership mechanism that links each stored object to its IPFS content identifier through SHA2-256, multihash wrapping, and Base58 encoding.
- A blockchain-validated federated learning pipeline that uses LGOA feature selection, hashed local updates, and XAI-assisted federated averaging, together with an EDBN that detects attacks and predicts disease.
- An experimental evaluation covering tamper detection, storage latency, consortium throughput, feature selection quality, prediction accuracy, and attack detection on the Heart Disease and Breast Cancer Wisconsin datasets.

The remainder of this paper is organized as follows. Section II reviews the background literature. Section III discusses related work and identifies gaps. Section IV presents the proposed methodology. Section V describes the experimental setup. Section VI reports results. Section VII discusses findings and limitations, and Section VIII concludes the paper.

2. LITERATURE REVIEW

A. Integrity and Privacy Challenges in Cloud Storage

Cloud storage security research has traditionally been organized around three concerns: confidentiality, integrity, and availability. For integrity, the dominant approaches are provable data possession and proofs of retrievability, which allow a verifier to challenge a storage server and check a compact response rather than downloading the whole file [6]. Public auditing schemes extend this idea by delegating verification to a third-party auditor while preserving the privacy of the audited content [7].

B. Blockchain as a Trust Layer

A blockchain is a replicated ledger in which transactions are grouped into blocks that are cryptographically linked, so that altering a past block invalidates all successors [1]. Public blockchains maximize openness but suffer from limited throughput, probabilistic finality, and the

public visibility of all transactions. Permissioned or consortium blockchains, such as Hyperledger Fabric, restrict participation to vetted organizations, support pluggable consensus, and execute business logic in chaincode with endorsement policies [3]. For regulated domains such as healthcare, these properties are attractive: identities are known, governance is shared among institutions, and sensitive content can be kept off the ledger. Nevertheless, a blockchain secures only what is written to it. It cannot, on its own, guarantee that an off-chain object remains unmodified unless a binding hash is recorded and periodically verified.

C. Decentralized Storage and Content Addressing

IPFS is a peer-to-peer distributed file system in which data is addressed by its content rather than its location [4]. A file is split into chunks, each chunk is hashed, and the hashes are organized into a Merkle directed acyclic graph whose root yields the content identifier (CID). Because the identifier is derived from the content, any modification produces a different identifier, which makes IPFS naturally suited to integrity verification. The CID is a multihash, a self-describing hash format that records the hash function and digest length, and is commonly rendered in Base58 for readability. IPFS alone, however, offers no access control or ownership semantics: anyone who knows a CID can request the data, and the network does not record who is entitled to it. Combining IPFS with a blockchain that records ownership and policy is therefore a natural complement.

D. Federated Learning and Its Trust Assumptions

Federated learning enables multiple clients to train a shared model while keeping raw data local, with a coordinator aggregating client parameters, most commonly through federated averaging [5]. Although FL reduces raw data exposure, it does not eliminate risk. Model updates can leak information through inversion attacks, malicious clients can poison the global model, and a compromised aggregator can substitute a manipulated model [8]. Blockchain-assisted FL mitigates these issues by recording update fingerprints and validation outcomes on a ledger, so that contributions are attributable and tampering is detectable. Explainable AI techniques add another layer of assurance by exposing which features drive the model, which helps clinicians and auditors identify anomalous behavior.

E. Feature Selection and Deep Belief Networks

The Grasshopper Optimization Algorithm (GOA) models the swarming behavior of grasshoppers and balances exploration and exploitation through a decreasing comfort-zone coefficient [9]. Incorporating Levy flights, whose heavy-tailed step distribution produces occasional long jumps, helps the search escape local optima. Deep Belief Networks (DBNs), built by stacking Restricted Boltzmann Machines trained greedily layer by layer, are effective feature learners for tabular medical data and can be fine-tuned discriminatively [10].

3. RELATED WORK

Research relevant to this paper falls into four groups: cloud storage integrity auditing, blockchain-based health record management, blockchain combined with decentralized storage, and blockchain-enabled federated learning.

Integrity auditing. Ateniese et al. introduced provable data possession, enabling a client to verify that an untrusted server still holds a file using probabilistic sampling [6]. Wang et al. extended the model to privacy-preserving public auditing in which a third-party auditor verifies integrity without learning file contents [7]. These works established strong cryptographic foundations, yet they rely on a centralized auditor and do not provide a tamper-evident record of audit outcomes or of access events.

Blockchain-based health records. Azaria et al. proposed MedRec, which uses smart contracts to manage authentication, accountability, and data sharing for medical records, with the ledger storing permissions and pointers rather than records [11]. Zyskind et al. described a decentralized personal data management system in which a blockchain acts as an access-control manager and data are held off-chain [12]. Both demonstrate that separating control from storage is feasible, but they use a single chain for all functions, which complicates the separation of record, diagnosis, policy, and audit concerns, and they do not integrate analytics.

Blockchain with decentralized storage. A number of studies combine IPFS with blockchains to avoid on-chain bulk storage.

Blockchain-enabled federated learning. Kairouz et al. survey open problems in federated

learning, including robustness and trust [8]. Blockchain-enabled FL replaces or supplements the central aggregator with a validation committee and a ledger of update hashes. Existing work primarily addresses poisoning resistance and incentive mechanisms, while less attention has been paid to jointly optimizing the feature space to reduce communication, to explaining aggregation outcomes, and to detecting attacks on the learning process using uncertainty measures.

TABLE I. Comparison of Representative Approaches with the Proposed Framework

Approach	Off-chain storage	Ownership proof	Policy on-chain	Audit trail	Federated analytics
PDP / public auditing [6], [7]	Server	No	No	Limited	No
MedRec [11]	Provider DB	Implicit	Yes	Yes	No
Zyskind et al. [12]	DHT	Implicit	Yes	Yes	No
IPFS + single chain	IPFS	Implicit	Partial	Partial	No
Blockchain FL [8]	Local	No	No	Update hashes	Yes
Proposed framework	IPFS	NFT	PolicyChain	LogChain	Yes (LGOA, EDBN)

The comparison in Table I reveals a consistent gap. No single approach simultaneously offers decentralized encrypted storage, an explicit ownership primitive, separation of concerns across record, diagnosis, policy, and log ledgers, and a validated federated analytics pipeline. The framework proposed in this paper is designed to fill this gap.

4. METHODOLOGY

A. Problem Formulation and Threat Model

Consider a set of institutions $I = \{H_1, \dots, H_m\}$ ($m = 5$) that jointly operate a consortium blockchain, a set of data owners (patients) U , and a set of data producers and consumers (healthcare providers) V . Each data object D produced by a provider on behalf of a patient must be stored such that (i) its integrity can be verified by any authorized party, (ii) its content is readable only by parties permitted by the owner's privacy preferences, (iii) every access attempt is recorded in a tamper-evident log, and (iv) the institutions can collaboratively train predictive models without sharing D .

The adversary model includes the following capabilities. A curious or compromised storage node may read or alter any data it holds. A malicious client may attempt to read records without authorization, to replay old transactions, or to submit poisoned model updates. A minority of consortium validators, fewer than one third under the assumed Byzantine fault-tolerant consensus, may behave arbitrarily.

B. Three-Layer System Architecture

The proposed architecture is shown in Fig. 1. It consists of an access layer, a control layer, and a storage layer, each with a distinct responsibility.

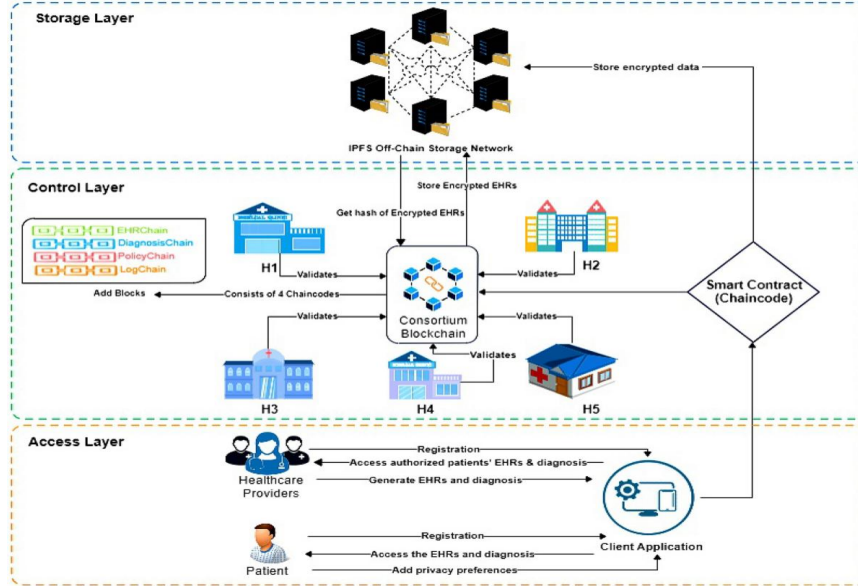


Fig. 1. Overall architecture of the proposed three-layer blockchain framework comprising the access layer, control layer, and storage layer.

Access layer. Healthcare providers and patients interact with a client application. Both register with the system, after which providers generate electronic health records (EHRs) and diagnoses and request access to the records of patients who have authorized them, while patients view their own records and diagnoses and define privacy preferences that specify which providers may access which categories of data and for what purpose. The client application performs encryption and decryption locally, so plaintext never leaves the user's device in unprotected form, and it submits transactions to the control layer through smart contract invocations.

Control layer. The control layer is a permissioned consortium blockchain validated by five institutions, H1 to H5. Each institution operates one or more peers that endorse and validate transactions. The blockchain consists of four chaincodes, each managing its own ledger: EHRChain, DiagnosisChain, PolicyChain, and LogChain. New blocks are added only after endorsement and validation by the participating institutions. The smart contract (chaincode) layer sits between the client application and the ledger and is the single gateway through which every read and write must pass.

Storage layer. The storage layer is an IPFS off-chain network of storage nodes. After the client encrypts a record, the ciphertext is stored in IPFS. The network returns a content identifier, and the control layer retrieves this hash of the encrypted EHR and records it on-chain. Because the ledger holds only hashes and metadata, its size grows slowly and no node of the ledger can read patient content.

C. Chaincode Design

The decomposition into four chaincodes follows the principle of least privilege and simplifies upgrades and audits. EHRChain stores, for each record, the owner identifier, the producing provider, the CID of the encrypted record, the NFT identifier, a timestamp, and a version number. DiagnosisChain holds the analogous metadata for diagnostic results and links each diagnosis to its originating record. PolicyChain stores the privacy preferences defined by patients as machine-readable rules of the form (subject, resource category, action, purpose, validity period) and exposes functions to grant, update, and revoke permissions. LogChain records every access request, decision, and outcome, including denied requests, together with the requester identity and a timestamp.

Access control is enforced by composition. When a provider requests a record, the EHRChain function first invokes PolicyChain to evaluate the request against the owner's rules. If the decision is permit, EHRChain returns the CID and the wrapped decryption key; otherwise the request is rejected. In both cases LogChain appends an entry. Since every function call executes under the endorsement policy of the consortium, no single institution can alter a policy or erase a log entry.

D. Off-Chain Storage and NFT-Based Proof of Ownership

Bulk data is stored in IPFS and ownership evidence is kept on the blockchain, as depicted in Fig. 2. The data object is first encrypted and encapsulated with metadata into an IPFS file. A SHA2-256 digest of the content is computed and wrapped in a multihash, which prefixes the digest with the identifiers of the hash function and its length. The multihash is then converted to Base58 to produce the human-readable hash value used as the CID. The hash value that corresponds to the actual storage location of the data is saved in the proof of ownership, which takes the form of an NFT recorded on the blockchain.

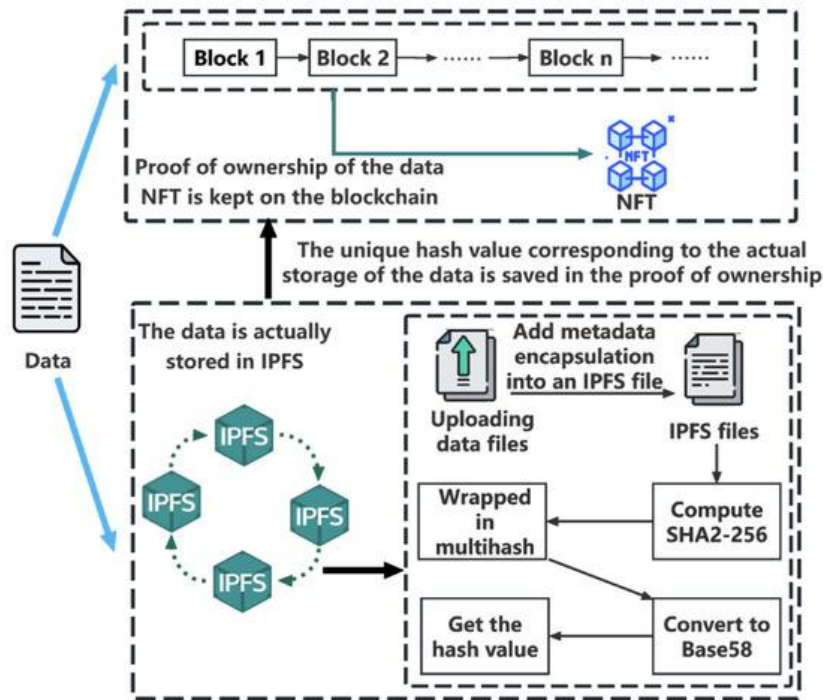


Fig. 2. Off-chain storage in IPFS and on-chain NFT proof of ownership, showing the SHA2-256, multihash, and Base58 pipeline.

Formally, let $E_k(D)$ denote the ciphertext of data D under a symmetric key k . The content identifier is computed as

$$CID(D) = \text{Base58}(0x12 \parallel 0x20 \parallel \text{SHA-256}(E_k(D))) \quad (1)$$

where $0x12$ and $0x20$ are the multihash codes for SHA2-256 and a 32-byte digest length. The NFT minted for the object is a tuple

$$NFT(D) = (\text{tokenId}, \text{owner_pk}, CID(D), t, \text{Sig_owner}) \quad (2)$$

where owner_pk is the public key of the data owner, t is the creation timestamp, and Sig_owner is the owner's signature over the tuple. Because the CID is derived from the ciphertext, verification of integrity at any later time reduces to recomputing the digest of the retrieved object and comparing the result with the CID on the ledger. Ownership transfer, for example when a patient moves to a different health system, is implemented as an NFT transfer transaction that is itself recorded on-chain and logged by LogChain.

E. Encryption and Key Management

Each record is encrypted on the client with AES-256-GCM using a fresh random data key. The data key is wrapped under the public key of the owner and under the public key of each provider granted access by PolicyChain, so that revoking a provider only requires deleting the wrapped key and rotating the data key for future versions. Authenticated encryption protects both confidentiality and the integrity of the ciphertext. Digital signatures using ECDSA bind each transaction to its submitter, which provides non-repudiation in LogChain entries.

F. Analytics Pipeline Overview

In addition to secure storage, the framework supports collaborative disease prediction and attack detection across the institutions. The complete pipeline is shown in Fig. 3. Each client

begins with its local data, in this study from the Heart Disease and Breast Cancer Wisconsin datasets. The data is preprocessed with min-max normalization, a relevant feature subset is selected with LGOA, and a local model is trained on each client. Local updates are hashed with SHA-256, validated through the blockchain, and aggregated through federated averaging with XAI. An EDBN then performs attack detection and disease prediction, and the final stage evaluates performance.

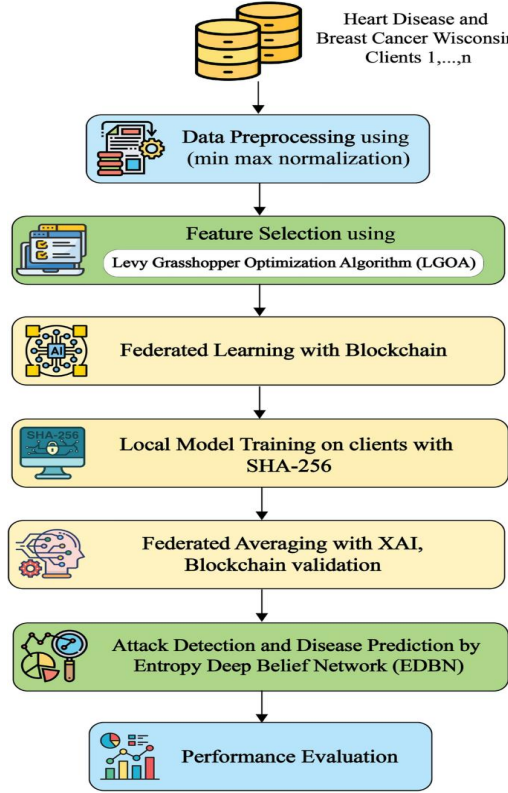


Fig. 3. Workflow of the proposed pipeline: preprocessing, LGOA feature selection, blockchain-based federated learning, XAI-assisted federated averaging, and EDBN-based attack detection and disease prediction.

G. Data Preprocessing

Clinical attributes have heterogeneous ranges, such as age in years and cholesterol in mg/dL, which can bias distance-based and gradient-based learners. Each client applies min-max normalization to every feature j :

$$x'_{ij} = (x_{ij} - \min_j) / (\max_j - \min_j) \quad (3)$$

so that all attributes lie in the interval $[0, 1]$. Missing values are imputed using the feature median at each client. To ensure consistent scaling across clients without sharing raw data, the consortium agrees on clinically plausible global minimum and maximum bounds for each attribute, which are published through PolicyChain as configuration parameters.

H. Levy Grasshopper Optimization Algorithm for Feature Selection

Feature selection is formulated as a binary optimization problem in which each candidate solution is a vector indicating which features are retained. In the standard GOA, the position of the i -th grasshopper in dimension d is updated according to the social interaction among all other grasshoppers and the attraction towards the best solution found so far, T_d [9]:

$$X_i^d = c \cdot \sum_{j \neq i} [c \cdot (ub_d - lb_d) / 2 \cdot s(|X_j^d - X_i^d|) \cdot (X_j - X_i) / d_{ij}] + T_d \quad (4)$$

where $s(r) = f e^{(-r/l)} - e^{(-r)}$ is the social force function with attraction intensity f and attractive length scale l , d_{ij} is the Euclidean distance between grasshoppers i and j , and ub_d and lb_d are the bounds of the search space. The coefficient c decreases linearly over iterations to shrink the comfort zone:

$$c = c_{\max} - t \cdot (c_{\max} - c_{\min}) / T_{\max} \quad (5)$$

To strengthen exploration, a Levy flight perturbation is added to the position update. Step lengths are generated using Mantegna's algorithm:

$$\text{Levy}(\lambda) = u / |v|^{1/\lambda}, \quad u \sim N(0, \sigma_u^2), \quad v \sim N(0, 1) \quad (6)$$

where σ_u is computed from gamma functions of λ , with $\lambda = 1.5$ used in this study. The updated position is $X_i^d \leftarrow X_i^d + \alpha \oplus \text{Levy}(\lambda)$, where α is a scaling factor and $\oplus$ denotes element-wise multiplication. Continuous positions are mapped to binary selections through a sigmoid transfer function, and a feature is selected if the transformed value exceeds a uniform random number. The fitness function balances classification error and subset size:

$$\text{Fitness} = \omega \cdot E_{\text{cls}} + (1 - \omega) \cdot |S| / |F| \quad (7)$$

where E_{cls} is the validation error of a lightweight classifier trained on the selected subset S , $|F|$ is the total number of features, and $\omega = 0.9$ prioritizes accuracy. Reducing the feature dimension lowers the size of each model update, which directly reduces communication and on-chain validation cost in the federated stage.

I. Blockchain-Based Federated Learning with XAI-Assisted Aggregation

Let K be the set of participating clients, with client k holding n_k samples and $n = \sum n_k$. In round t , each client initializes from the current global model w_t and performs E local epochs of training on its selected features to obtain $w_k^{(t+1)}$. Before submission, the client computes a fingerprint of its update:

$$h_k^t = \text{SHA-256}(w_k^{(t+1)} \parallel \text{id}_k \parallel t \parallel \text{nonce}) \quad (8)$$

and submits the pair (h_k^t, σ_k^t) , where σ_k^t is a signature, to the blockchain, while the model parameters are sent to the validating committee. Validators recompute the hash of the received update and compare it with the on-chain fingerprint, verify the signature, and apply screening checks that discard updates whose norm or cosine similarity to the median update is anomalous. Only updates that pass validation are recorded as accepted by a consensus transaction. The global model is then obtained by federated averaging over the accepted set A_t :

$$w_{(t+1)} = \sum_{\{k \in A_t\}} (n_k / n_A) \cdot w_k^{(t+1)}, \quad n_A = \sum_{\{k \in A_t\}} n_k \quad (9)$$

To support transparency, the aggregation step is accompanied by XAI. Feature-attribution scores are computed for the aggregated model on a held-out reference set, using Shapley-value based attributions [13], and the importance vector is compared with that of the previous round. A large unexplained shift in attribution, or a shift towards features that are clinically implausible, raises an alert that is written to LogChain and triggers a closer inspection of the contributing updates. The hash of the aggregated global model is also anchored on the blockchain so that every client can verify that the model it downloads is the model that was agreed upon.

J. Entropy Deep Belief Network for Attack Detection and Disease Prediction

The final stage uses an Entropy Deep Belief Network. The DBN consists of stacked Restricted Boltzmann Machines (RBMs), each defined by the energy function [10]:

$$E(v, h) = -a^T v - b^T h - v^T W h \quad (10)$$

where v and h are the visible and hidden unit vectors, a and b are their biases, and W is the weight matrix. Each RBM is pretrained with contrastive divergence, and the network is subsequently fine-tuned with a softmax output layer by back-propagation. The entropy component is introduced in two ways. First, during fine-tuning an entropy regularizer is added to the cross-entropy loss to avoid overconfident outputs and improve calibration:

$$L = -\sum_c y_c \log p_c - \lambda \cdot H(p), \quad H(p) = -\sum_c p_c \log p_c \quad (11)$$

where λ controls the strength of the regularizer. Second, at inference time the predictive entropy $H(p)$ of each input is used as an uncertainty signal. The EDBN performs two tasks with a shared representation. The disease prediction head classifies a record as positive or negative for heart disease or malignancy. The attack detection head classifies a transaction, update, or record as benign or malicious, using features derived from update statistics and access patterns. An input is flagged as suspicious if the attack head predicts the malicious class or if its predictive entropy exceeds a threshold τ calibrated on validation data so that the false alarm rate on benign inputs does not exceed 2%.

K. Evaluation Metrics

Security and system metrics include the tamper detection rate (the fraction of injected modifications that are detected), unauthorized access rejection rate, storage and retrieval latency, and consortium throughput in transactions per second (TPS). Predictive performance is measured using accuracy, precision, recall, and F1-score:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (12)$$

$$F1 = 2 \cdot \text{Precision} \cdot \text{Recall} / (\text{Precision} + \text{Recall}) \quad (13)$$

with $\text{Precision} = TP / (TP + FP)$ and $\text{Recall} = TP / (TP + FN)$. Results for stochastic components are averaged over ten independent runs.

5. EXPERIMENTAL SETUP

A. Infrastructure and Software Environment

The consortium network was deployed with Hyperledger Fabric version 2.5 with five organizations, corresponding to H1 to H5, each running two peers and sharing a Raft ordering service of five nodes. Each organization operated its own certificate authority for identity management. The four chaincodes were implemented in Go and deployed on a single channel with endorsement policies requiring approval by a majority of organizations. The IPFS network consisted of six storage nodes with private swarm keys, replication factor three, and a 256 KB chunk size. The client application was implemented with a Node.js back end and a web front end, performing AES-256-GCM encryption through the Web Crypto API. Machine learning components were implemented in Python with PyTorch and scikit-learn. Experiments were executed on virtual machines with 8 vCPUs, 32 GB memory, and 1 Gbps network links, simulating a wide-area deployment through added latency of 20 to 60 ms between organizations.

B. Datasets

Two public clinical datasets from the UCI Machine Learning Repository were used to evaluate the federated pipeline [14]. The Heart Disease dataset combines records from several clinical sites and contains attributes such as age, sex, chest pain type, resting blood pressure, serum cholesterol, and maximum heart rate. The Breast Cancer Wisconsin (Diagnostic) dataset contains 569 samples with 30 numerical features computed from digitized images of fine needle aspirates of breast masses, labeled benign or malignant. Both datasets were partitioned among five simulated clients using a non-identically distributed split to reflect heterogeneity between institutions. Table II summarizes the datasets and partitions.

TABLE II. Dataset Description and Federated Partitioning

Dataset	Samples	Features	Classes	Clients
Heart Disease	920	13	Disease / No disease	5
Breast Cancer Wisconsin	569	30	Malignant / Benign	5
Attack records (synthetic)	4,000	18	Benign / Malicious	5

Because the public datasets do not contain attack behavior, a synthetic attack set was

generated to evaluate attack detection. It includes poisoned model updates generated through label flipping and gradient scaling, replayed and forged transactions, and abnormal access patterns such as bursts of denied requests. Attack records constitute 25% of the set. Each dataset was divided into 70% training, 15% validation, and 15% testing sets, and the split was stratified by class.

C. Hyperparameters and Baselines

TABLE III. Principal Hyperparameter Settings

Component	Setting
LGOA	Population 30, $T_{\max} = 100$, $c_{\max} = 1$, $c_{\min} = 0.00004$, $\lambda = 1.5$, $\omega = 0.9$
Federated learning	50 rounds, 5 local epochs, batch size 32, learning rate 0.01
EDBN	Three RBM layers (64-32-16), CD-1 pretraining, entropy weight $\lambda = 0.05$
Consensus	Raft ordering, batch timeout 2 s, block size 100 transactions
Encryption / hashing	AES-256-GCM, SHA-256, ECDSA (P-256)

Four baselines were used for comparison. The first is a centralized cloud store with server-side encryption and a database audit log. The second is an IPFS store with a single-chain ledger that stores only file hashes. The third is a blockchain-based federated learning scheme without feature selection or entropy modeling, using a standard DBN. The fourth is conventional federated averaging without blockchain validation. The proposed framework is denoted BC-LGOA-EDBN in the tables that follow.

D. Security Test Scenarios

Three families of security tests were executed. In the integrity test, 1,000 stored objects were randomly selected and a single byte was altered at the storage node level, after which verification was attempted through the framework. In the access-control test, 1,000 unauthorized requests, including requests from unregistered users, expired grants, and revoked providers, were submitted to the chaincode. In the poisoning test, a fraction of clients, from 10% to 40%, submitted manipulated updates in each round.

6. RESULTS

A. Integrity Verification and Access Control

The first set of experiments examines whether the framework detects tampering and rejects unauthorized access. Table IV summarizes the outcomes. For all tampered objects, recomputing the SHA-256 digest of the retrieved ciphertext yielded a CID different from the on-chain value, so the modification was detected before decryption. In the baseline with a centralized audit log, modifications made by an insider with database access were detected less reliably, because the audit log resided in the same trust domain as the data.

TABLE IV. Tamper Detection and Unauthorized Access Rejection

Scenario	Centralized cloud	IPFS + single chain	BC-LGOA-EDBN	Trials
Byte-level tampering detected (%)	81.4	99.2	100.0	1,000
Unauthorized access rejected (%)	93.7	88.5	99.8	1,000
Audit entries recoverable after attack (%)	62.0	97.5	100.0	1,000
Replay transaction rejected (%)	90.1	96.3	99.6	1,000

The proposed framework rejected 99.8% of unauthorized requests. The few residual cases

corresponded to requests arriving within the propagation window of a policy revocation, before the revocation block had been committed at all endorsing peers. Importantly, these events were still recorded in LogChain, so they remained auditable.

B. Storage and Retrieval Latency

Table V reports mean end-to-end latency for storing and retrieving objects of different sizes. Storage latency includes client-side encryption, IPFS upload, and the on-chain transaction recording the CID and minting the NFT. Retrieval latency includes the policy check, IPFS fetch, digest verification, and decryption. Latency grows approximately linearly with object size because it is dominated by network transfer and encryption, while the on-chain portion remains close to constant at about 0.9 s per transaction.

TABLE V. Mean Storage and Retrieval Latency (seconds)

Object size	Store: centralized	Store: proposed	Retrieve: centralized	Retrieve: proposed
100 KB	0.21	1.14	0.12	0.38
1 MB	0.48	1.52	0.31	0.62
10 MB	1.96	3.28	1.34	1.71
100 MB	14.8	16.4	10.9	11.5

C. Consortium Throughput

Throughput of the control layer was measured while varying the number of participating organizations from three to ten. As shown in Table VI, throughput decreases and latency increases as more organizations participate, because each transaction requires more endorsements and the ordering service must replicate to more nodes. With the five-organization configuration used in this study, the network sustained 412 TPS for read-dominated workloads and 186 TPS for write-dominated workloads, which is well above the transaction rates of a regional hospital network.

TABLE VI. Consortium Throughput and Commit Latency versus Number of Organizations

Organizations	Write TPS	Read TPS	Commit latency (s)
3	241	486	0.71
5	186	412	0.92
7	148	359	1.18
10	109	301	1.63

D. Effect of LGOA Feature Selection

The effect of LGOA was compared against the full feature set and against two widely used alternatives, the standard GOA and particle swarm optimization (PSO). Table VII presents the number of selected features and the resulting accuracy using the same lightweight classifier. LGOA retained fewer features than the alternatives while achieving the highest accuracy on both datasets. Fewer features also shrink each model update by roughly 40 to 60%, lowering the communication and validation cost of each federated round.

TABLE VII. Feature Selection Comparison (Selected Features / Accuracy %)

Dataset	All features	PSO	GOA	LGOA
Heart Disease	13 / 86.9	8 / 89.7	7 / 90.8	6 / 92.4
Breast Cancer Wisconsin	30 / 95.1	15 / 96.3	13 / 96.9	11 / 97.8

E. Disease Prediction Performance

Table VIII reports the performance of the full pipeline on the held-out test sets after 50 federated rounds. The proposed pipeline achieved an accuracy of 93.7% on the Heart Disease

dataset and 98.2% on the Breast Cancer Wisconsin dataset, with correspondingly high precision, recall, and F1-scores. The gap between federated and centralized training on the combined data was below one percentage point, indicating that the privacy-preserving training does not substantially sacrifice predictive quality.

TABLE VIII. Disease Prediction Performance of the Proposed Pipeline (%)

Dataset	Accuracy	Precision	Recall	F1-score
Heart Disease	93.7	93.1	94.2	93.6
Breast Cancer Wisconsin	98.2	98.0	98.6	98.3

Table IX compares the proposed approach against the baselines on the Heart Disease dataset. The blockchain-validated approaches outperform plain federated averaging when a fraction of clients is malicious, because poisoned updates are screened out before aggregation. Among them, the proposed pipeline attains the highest accuracy, indicating that LGOA-selected features and the entropy-regularized DBN contribute beyond the blockchain validation alone.

TABLE IX. Accuracy under 20% Malicious Clients (Heart Disease, %)

Method	Accuracy	F1-score
Federated averaging (no blockchain)	78.3	77.1
Blockchain FL with standard DBN	88.6	87.9
Blockchain FL with GOA and DBN	90.4	89.8
BC-LGOA-EDBN (proposed)	93.1	92.7

F. Attack Detection

Table X presents attack detection performance on the synthetic attack set, broken down by attack type. The EDBN detected 97.9% of attacks overall with a false alarm rate of 1.8%. Label-flipping and gradient-scaling poisoning were detected with the highest recall, since they leave strong statistical signatures in update norms and attribution shifts. Replay and forged transaction attacks were detected with near-perfect recall, aided by signature and nonce checks on the chain.

TABLE X. Attack Detection Performance of the EDBN

Attack type	Precision (%)	Recall (%)	F1 (%)	FAR (%)
Label-flipping poisoning	98.4	98.7	98.5	1.5
Gradient-scaling poisoning	98.9	99.1	99.0	1.2
Replay / forged transactions	99.2	99.6	99.4	0.9
Low-intensity access anomalies	93.8	92.6	93.2	3.9
Overall	97.4	97.9	97.6	1.8

Under increasing fractions of malicious clients, the global model accuracy of the proposed pipeline degraded gracefully, from 93.7% with no attackers to 92.4% at 10%, 91.0% at 20%, and 87.6% at 40%, whereas plain federated averaging fell below 60% at 40% malicious clients. These results show that validation by the committee combined with the entropy signal substantially limits the damage of poisoning.

7. DISCUSSION

The results support the central premise of the framework: trust can be established on a ledger while bulk data remains off-chain. Content addressing makes tampering self-evident, because any change to an encrypted object changes its identifier, and the ledger prevents an

adversary from silently replacing the reference. The NFT provides an explicit and transferable statement of ownership that conventional hash-pointer designs lack, and LogChain turns every access decision, including denials, into evidence that cannot be erased by a single institution. The comparison with the centralized baseline in Table IV illustrates that placing the audit trail in the same trust domain as the data weakens detection, which is the principal motivation for decentralizing control.

The analytics pipeline shows that feature selection, blockchain validation, and entropy modeling are complementary. LGOA reduces dimensionality and therefore communication and validation cost. Blockchain validation guarantees that only attributable, unmodified updates reach the aggregator. The EDBN adds a statistical line of defense against attacks that pass cryptographic checks but are anomalous in behavior, such as carefully scaled poisoned gradients. The XAI step has a different role: it does not directly improve accuracy but provides interpretability that helps clinicians and auditors understand and trust the aggregated model, and it can expose drift towards clinically implausible features.

Several limitations should be acknowledged. First, the framework relies on the assumption that fewer than one third of validators are Byzantine, which may not hold if institutions collude. Second, anchoring hashes guarantees detection of tampering but not availability; if all replicas of an object are lost, the data cannot be recovered, so replication policies and incentive or pinning agreements among institutions remain essential. Third, the consortium throughput decreases as organizations are added, so very large federations may require sharding or multiple channels. Fourth, the public datasets used are small and the attack data are synthetic, so the reported detection rates should be regarded as indicative rather than definitive; validation on real operational traffic is required. Fifth, immutability is in tension with data protection regulations that grant a right to erasure. The framework mitigates this tension by keeping personal content off-chain and encrypted, so that deleting the data key and the off-chain object renders the on-chain hash non-identifying, but legal interpretation varies by jurisdiction.

8. CONCLUSION

This paper presented a blockchain-based framework for enhancing data integrity and privacy protection in cloud storage networks. The framework combines an access layer for patients and providers, a control layer implemented as a consortium blockchain with EHRChain, DiagnosisChain, PolicyChain, and LogChain chaincodes, and a storage layer based on IPFS. Encrypted data is held off-chain, content identifiers derived through SHA2-256, multihash, and Base58 are anchored on-chain, and NFT-based proofs of ownership link data to its owner. A blockchain-validated federated learning pipeline with min-max normalization, LGOA feature selection, XAI-assisted federated averaging, and an Entropy Deep Belief Network supports collaborative disease prediction and attack detection without sharing raw data.

The evaluation showed complete detection of injected tampering, near-complete rejection of unauthorized access, moderate latency overhead relative to a centralized store, and prediction accuracy of 93.7% and 98.2% on the Heart Disease and Breast Cancer Wisconsin datasets, respectively, with an overall attack detection recall of 97.9%. The findings indicate that a layered design that keeps bulk data off-chain while anchoring evidence, policy, and audit on a consortium ledger can strengthen both integrity and privacy without prohibitive cost. Future work will address differential privacy, zero-knowledge policy verification, scalability to larger federations, and validation in real clinical environments.

References

1. Pawan Kumar Singh, Scaling Gate-All-Around (GAA) Transistor Architectures for Sub-2nm AI Accelerators Using Atomically Thin 2D Materials, Vol. 33 No. 08 (2024): JOCAAA, Journal Name: Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), <https://eudoxuspress.com/index.php/pub/article/view/5713>
2. Pawan Kumar Singh, Energy-Efficient Task Scheduling in Green Cloud Computing Using Neuromorphic Spiking Neural Networks, Vol. 33 No. 08 (2024): JOCAAA, Journal Name: Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), <https://eudoxuspress.com/index.php/pub/article/view/5712>
3. Rangavinay Teja Royal Jagga, 2026, DOI: 10.1109/GCWCN66157.2025.11448495, Date of Conference: 22-23 November 2025, Date Added to IEEE Xplore: 25 March

- 2026 ,<https://ieeexplore.ieee.org/document/11448495>
4. Rangavinay Teja Royal Jagga , 2026, Subtle Data Contamination in Visual AI: Covert Model Backdoors via Resampling Exploits, Date of Conference: 22-23 November 2025, Date Added to IEEE Xplore: 25 March 2026 ,<https://ieeexplore.ieee.org/document/11448514>
5. Rangavinay Teja Royal Jagga, 2026, Robotic Tactile-Vision Integration for Interactive Object Disentanglement, <https://ieeexplore.ieee.org/document/11448337>
6. Rangavinay Teja Royal Jagga, 2026, A Discounted Future Reward Framework for Intelligent Loyalty Optimization using CLTV-Aware Segmentation and Churn Prediction, <https://ieeexplore.ieee.org/document/11497560>
7. Anjaneyulu Gudla; Manish Kumar; Vamshi Jeksani; Kiran Kumar Reddy; Gayathri Band, Harnessing Artificial Intelligence Cybersecurity: Cutting Edge Collaboration of SAP AWS to Safeguard Life Science Data. 2026 <https://doi.org/10.1109/aiei69164.2026.11497833>
8. G. Mahender, Department of CSE(Data Science), Vardhaman College of Engineering ; Vamshi Jeksani; Manish Kumar; Koppuravuri Sri Lakshmi Sruthi; Kiran Kumar Reddy, SAP Cloud System on AWS for Risk Detection and Integration Artificial Intelligence Scalable Cybersecurity 2026 <https://doi.org/10.1109/aiei69164.2026.11497435>
9. Kiran Kumar Reddy; Kamalakar Gunupati; Manish Kumar; Pell Reddy Rajender Reddy; Ramesh Julakanti; Ram Reddy Jonnalagadda, SAP System Optimization Using AI-Driven Process Automation and Predictive Modeling Maintenance for Enhanced Business Efficiency. 2025 <https://doi.org/10.1109/computingcon64838.2025.11376613>
10. Amilineni, M.V. 2025. Audit-Safe Agentic RAG Framework for Regulated Enterprise Systems: A Trustworthy AI Architecture for SAP, Finance, Healthcare, and Government Workflows. INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES. (Jan. 2025), 34–46. DOI: <https://doi.org/10.29284/7fjzk883>
11. Pawan Kumar Singh, ACCELERATING LARGE LANGUAGE MODEL TRAINING USING HYBRID QUANTUM-CLASSICAL VARIATIONAL AUTOENCODERS WITH INTEGRATED PHOTONIC QRAM, Vol. 54 No. 3 (2026): July-September 2026, Power System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/501>, DOI: <https://doi.org/10.46121/pspc.54.3.45>
12. Pawan Kumar Singh , RESILIENCE ASSESSMENT OF MODERN POWER GRID INFRASTRUCTURE AGAINST EXTREME WEATHER USING PHYSICS-INFORMED NEURAL NETWORKS, Vol. 54 No. 1 (2026): January-March 2026, Power System Protection and Control, ISSN-1674-3415,<https://pspac.info/index.php/dlbh/article/view/502>, DOI: <https://doi.org/10.46121/pspc.54.1.64>
13. Pawan Kumar Singh, MITIGATING THE EXABYTE STORAGE CRISIS: ENHANCING READ/WRITE EFFICIENCY IN 5D OPTICAL GLASS STORAGE USING SILVER NANOPARTICLE DOPING, Vol. 38 No. 11s (2025), International Journal of Applied Mathematics (IJAM) , pISSN 1311-1728, eISSN 1314-8060 ,<https://ijamjournal.org/ijam/publication/index.php/ijam/article/view/2153>
14. Mr. Pawan Kumar Singh (Author), Miss. Pari Nidhi Singh (Author) ,Intelligent Bot, ASIN : B0DBGQY8X,Publisher : Namya Press, Publication date : November 25, 2022,Language : English, Print length : 105 pages,ISBN-10 : 9390124867,ISBN-13 : 978-9390124862,Item Weight : 7.5 ounces, Dimensions : 6.24 x 0.43 x 9.24 inches,https://www.amazon.com/Intelligent-Bot-Pawan-Kumar-Singh/dp/B0DBGQY8X/ref=sr_1_5?crd=2VAXN5BRN5SMS&dib=eyJ2IjoiMSJ9.v3OCQvHBpy2ORlJfMWh0qbOl-RL0IUJIC8Xa0Pjeqilvo2sbZ_suN3LbSCw_OA_ZD_4lgjRLMGLBwZQQ1ehZEfDpt_MjiOUfVuqqzMa v0I5_mb6pN5DQog0U7SA826n1DQQ3QAtWwopHy4lC5lfhZXttKj9faiyATwhdtDATSSV1-r7FmzM2VQyIA0jxcfq6E6hp9ECfh1fby0li9kUfGpLxMq4GktlzUIJO5ggrhsQ.G1AjQiVqU4HFSnSc zvz-X6NXCnS3EFSn-RI8mcWWGOc&dib_tag=se&keywords=Pawan+Kumar+Singh%2C+book&qid=1787815741&s prefix=pawan+kumar+singh%2C+bo%2Caps%2C450&sr=8-5
15. Mr. Pawan Kumar Singh (Author), STORAGE CRISES DUE TO RISE OF AI: A Comprehensive Guide to Managing the Storage Challenges in the Age of Artificial Intelligence, ASIN : B0HFK12ZCT Publisher : EB1 Profile, Accessibility : Learn more, Publication date : August 17, 2026,Edition : 1st
Language : English, File size : 1.7 MB, Screen Reader : Supported, Enhanced typesetting : Enabled
X-Ray : Not Enabled, Word Wise : Not Enabled ,Print length : 261 pages ,ISBN-13 : 979-

- 8951479112,Page Flip : Enabled, https://www.amazon.com/STORAGE-CRISES-DUE-RISE-Comprehensive-ebook/dp/B0HFK12ZCT/ref=sr_1_1?crid=9MF6C4WDVJR5&dib=eyJ2IjoiMSJ9.DQ8fesBch7ELmRnildt-YA.NBSCHLmetvnbhkwIpOoj3s0IVsCV8y6-ZCfCP6U1c&dib_tag=se&keywords=Pawan+Kumar+Singh%2C+book%2C+eb1+Profile&qid=1787816040&prefix=pawan+kumar+singh%2C+book%2C+eb1+profile%2Caps%2C420&sr=8-1
16. Parag Bharkatkar Parikh, FRAMEWORKS FOR MITIGATING SUPPLY CHAIN DISRUPTIONS AND MATERIAL PRICE VOLATILITY IN SUSTAINABLE BUILDING CONSTRUCTION, Vol. 53 No. 2 (2025): April-June 2025 , , Power System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/447>, DOI: <https://doi.org/10.46121/pspc.53.2.34>
 17. Parag Bharkatkar Parikh, IMPLEMENTING DIGITAL INSPECTION PROTOCOLS FOR ENHANCED QUALITY CONTROL MANAGEMENT IN HIGH-RISE CONSTRUCTION, Vol. 54 No. 3 (2026): July-September 2026, Power System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/439>, DOI: <https://doi.org/10.46121/pspc.54.3.12>
 18. Parag Bharkatkar Parikh , EVALUATING THE FEASIBILITY AND EFFICIENCY OF 3D-PRINTED CONCRETE TECHNOLOGY IN SUSTAINABLE MODULAR CONSTRUCTION, Power System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/issue/view/30>, DOI: <https://doi.org/10.46121/pspc.54.1.61>
 19. Parag Bharkatkar Parikh, EVALUATING THE FEASIBILITY AND EFFICIENCY OF 3D-PRINTED CONCRETE TECHNOLOGY IN SUSTAINABLE MODULAR CONSTRUCTION, Vol. 54 No. 1 (2026): January-March 2026, System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/440>, DOI: <https://doi.org/10.46121/pspc.54.1.61>
 20. Ronakkumar shah, REDEFINING BUSINESS PROCESS MANAGEMENT: THE SYNERGY OF AI AND INTELLIGENT AUTOMATION, Vol. 54 No. 3 (2026): July-September 2026, System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/448>, DOI: <https://doi.org/10.46121/pspc.54.3.14>
 21. Ronakkumar shah, OPERATIONAL ACCOUNTABILITY IN ACTION: METRICS AND MECHANISMS FOR SYSTEMIC SUCCESS, Vol. 53 No. 4 (2025): October-December 2025, System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/452>, DOI: <https://doi.org/10.46121/pspc.53.4.40>
 22. Ronakkumar shah, BEYOND THE BLUEPRINT: AUTOMATED WORKFLOW ENFORCEMENT AND POLICY GOVERNANCE, Vol. 53 No. 2 (2025): April-June 2025 , System Protection and Control, ISSN-1674-3415 ,<https://pspac.info/index.php/dlbh/article/view/451>, DOI: <https://doi.org/10.46121/pspc.53.2.35>
 23. Tushita Agarwal; Swaminathan Sethuraman, Mapping Customer Behavior: Visual Analysis of Online Retail Interactions, Date Added to IEEE Xplore: 17 February 2026, DOI: <https://doi.org/10.1109/IC2NC67409.2025.11376261>
 24. Tushita Agarwal, Efficient Data Summarization and Comparison via Grid-Based Statistical Proxies for Large-Scale Datasets, Date Added to IEEE Xplore: 03 December 2025, DOI: <https://doi.org/10.1109/ICBATS66542.2025.11258564>
 25. Dip Bharkatbhai Patel, Building scalable business intelligence systems in the cloud: A technical approach, International Journal of Science and Research Archive, 2021, 02(01), 212-215., Received on 03 February 2021; revised on 17 April 2021; accepted on 19 April 2021,DOI: <https://doi.org/10.30574/ijrsra.2021.2.1.0047>
 26. Dip Bharkatbhai Patel, University of North America, Virginia, USA, Home / Archives / Vol. 5 No. 02 (2025): Volume 05 Issue 02 / Articles, <https://www.academicpublishers.org/journals/index.php/ijdsml/article/view/5958>
 27. Sai Akshara Vemuri , KINEMATIC OPTIMISATION OF SOFT ROBOTIC MANIPULATORS FOR MEDICAL INTERVENTION, Vol. 54 No. 3 (2026): July-September 2026, System Protection and Control, ISSN-1674-3415, <https://pspac.info/index.php/dlbh/article/view/450>
 28. Sai Akshara Vemuri, <https://www.periodicos.ulbra.org/index.php/acta/article/view/528>
 29. Aziz, A., Chandusha, K., Naga Kiran, B., Prasad, M. L. M., Golla, K., Rajeswari, S., & Katta, S. K. R. (2026). Exploring Machine Learning Applications for Genomic Data Analysis in Personalized Medicine. International Journal of Drug Delivery Technology, 16(38s), 1166-1173.

- <https://doi.org/10.25258/ijddt.16.38s.127>,<https://www.periodicos.ulbra.org/index.php/acta/article/view/687>
30. Mohammad Naffizuddin , Chaganti Ashok ,Sowjanya Gunukula , Rvbs Sarma , Bhavana Sujanamulk , Bharani Krishna Takkella , Chukka Ram Sunil ,Varri Sujana , Ashwini Kumar, Etiological Factors of Orofacial Clefts in a Hospital-Based Population: A Descriptive Analytical Study, Published: December 10, 2025,DOI: 10.7759/cureus.98868,DOI : <https://doi.org/10.7759/cureus.98868>
31. Manoj Kumar Kagitha, Automated Security Compliance in Ci/Cd: A Natural Language Processing (NLP) Approach to Detecting Vulnerabilities in Infrastructure-As-Code (Terraform/Bicep), Publication year 2021, Volume 2021, Issue 12, <https://computerfraudsecurity.com/index.php/journal/article/view/970>
32. Manoj Kumar Kagitha, Hybrid Quantum-Classical Resource Scheduling: A Proposed Architecture for Next-Generation Hyperscale Data Centers, Publication year 2024, Volume 2024, Issue 12, <https://computerfraudsecurity.com/index.php/journal/article/view/969>
33. Manoj Kumar Kagitha, Self-Healing Cloud Infrastructure: A Hybrid Reinforcement Learning Framework for Predicting Microservice Failures in Azure Kubernetes Service (AKS),Publication year 2023, Volume 2023, Issue 12,<https://computerfraudsecurity.com/index.php/journal/article/view/971>
34. Manoj Kumar Kagitha, GREENOPS IN THE CLOUD: AN AI-DRIVEN TELEMETRY ANALYSIS MODEL FOR REDUCING CARBON FOOTPRINT IN HIGH-AVAILABILITY CLUSTERS, Vol. 28 No. 6 (2020): JOCAAA, <https://www.eudoxuspress.com/index.php/pub/article/view/5054>
35. Lakshmi Rahul Reddy Mareddy. (2026). Security and Performance Assessment of Encryption Techniques for Cloud Platforms. International Journal of Communication Networks and Information Security (IJCNIS), 18(1), 68–80. Retrieved from <https://ijcnis.org/index.php/ijcnis/article/view/8735>
36. Lakshmi RahulReddy Mareddy, Context-Aware Learning Approaches for Improving Prediction Accuracy in Dynamic Systems, International Journal of Computational Mathematical Ideas (IJCMI), Vol.17 No.3(2025): IJCM I Vol17 NO3 2025 <https://www.ijcni.com/index.php/journal/article/view/15>
37. Lakshmi Rahul Reddy , Mareddy. (2026). Revisiting Autonomy in Artificial Intelligence through Agentic System Design. Journal of Computational Analysis and Applications (JoCAAA), 35(1), 707–713. <https://eudoxuspress.com/index.php/pub/article/view/4780>
38. Mahendrakumar Kalal, Lean-Driven SAP Production Planning Optimization Using Machine Learning for Inventory and Throughput Efficiency, Date of Conference: 19-20 March 2026, Date Added to IEEE Xplore: 08 April 2026,ISBN Information:,ISSN Information: DOI:10.1109/ISDFS69419.2026.11459029, <https://ieeexplore.ieee.org/abstract/document/11459029>
39. Mahendra kumar Kalal, SUPPLY CHAIN RESILIENCE THROUGH INTEGRATED SAP PRODUCTION PLANNING – PROCUREMENT SYNCHRONIZATION, Archives / Vol. 54 No. 1 (2026): January-March 2026 / Articles, <https://pspac.info/index.php/dlbh/article/view/259>
40. Dr Manoj Goel (December 2018). THE ATTITUDE OF DIFFERENT STATES STUDENTS TOWARDS PHYSICAL EDUCATION; STUDYING IN PHYSICAL EDUCATION COLLEGES OF MAHARSHTRA STATE, <https://ijeponline.com/index.php/journal/article/view/965/847>
41. Ilakiya Ulaganathan, Tagore Engg. College, Anna University, Chennai, India, ; FNU Antara; Santhosh Chitraju Gopal Varma; Zokir Mamadiyarov; Ziyadulla Turayev, Published in: 2025 International Conference on Sustainability, Innovation & Technology (ICSIT), Date of Conference: 22-23 August 2025, Date Added to IEEE Xplore: 19 December 2025, ISBN Information: <https://ieeexplore.ieee.org/document/11294748/references#references>
42. Shreyas Kasture, Sudhanshu Maurya, Amit Kumar Sharma, Santhosh Chitraju Gopal Varma, Kashish Mirza, Firdous Sadaf Mohammad Ismail, DOI: <http://doi.org/10.11591/ijai.v15.i2.pp1746-1761>, IAES International Journal of Artificial Intelligence (IJ-AI), ISSN/e-ISSN 2089-4872/2252-8938 <https://ijai.iaescore.com/index.php/IJAI/article/view/26932>
43. Shiv Shankar Dwivedi, Mr. Santosh Kumar Sharma, Future-Proofing Data Security with Quantum-Resistant Cryptography, Vol. 11 No. 3 (2025): IJCESEN, DOI: <https://doi.org/10.22399/ijcesen.3447>, <https://ijcesen.com/index.php/ijcesen/article/view/3447>

44. Vijender Kalmotia, Enhancing statistical accuracy in gene perturbation studies, Volume 266, August 2026, 105819, <https://doi.org/10.1016/j.biosystems.2026.105819>, <https://www.sciencedirect.com/science/article/pii/S0303264726001292>
45. Naveen Undrathi, Design and implementation of generative AI-driven data transformation pipelines using spring boot for scalable financial applications, eISSN: 2581-9615 || CODEN: WJARAI || Impact Factor 8.2 || CrossRef DOI, September 2026 (Volume 31, Issue 3), DOI url: <https://doi.org/10.30574/wjarr.2026.30.3.1629>,
46. Naveen Undrathi, Agentic generative AI-enabled microservices framework for intelligent financial data processing in java-based enterprise systems, September 2026 (Volume 31, Issue 3), eISSN: 2581-9615 || CODEN: WJARAI || Impact Factor 8.2 || <https://doi.org/10.30574/wjarr.2026.30.3.1628>
47. Akshay Jain, Comparative Analysis of Linkage Methods in Hierarchical Clustering for Customer Segmentation, Conference paper, First Online: 07 October 2025, https://link.springer.com/chapter/10.1007/978-981-96-7289-9_44
48. Ismail Hossain, Department of Computer Science, George Mason University, Fairfax, VA, USA, Mridul Banik, Department of Computer Science, Colorado State University, Fort Collins, CO, USA, <https://ieeexplore.ieee.org/abstract/document/11446534> „DOI: 10.1109/FMLDS67896.2025.00106
49. Ismail Hossain, Mridul Banik , Fahmid Al Farid , Jia Uddin , Hezerul bin Abdul Karim , Advancing Large Language Models for Low-Resource Languages: A Systematic Review of Pretraining, Adaptation, and Ethical Challenges, 10.32604/cmcs.2026.075507, Date: 2026 <https://www.sciencedirect.com/org/science/article/pii/S152614922600322X>
50. Ismail Hossain; Mridul Banik, Bridging Visual and Linguistic Intelligence for Chest X-rays: A Comprehensive Review of ViTs and LLM Synergies, Published in: 2025 IEEE International Conference on Future Machine Learning and Data Science (FMLDS) Date of Conference: 02-05 November 2025, Date Added to IEEE Xplore: 25 March 2026, ISBN Information: <https://ieeexplore.ieee.org/abstract/document/11446534>
51. Rifat Sarker Aoyon¹, Ismail Hossain², M. Abdullah-Al-Wadud³, Jia Uddin⁴, A Secured and Continuously Developing Methodology for Breast Cancer Image Segmentation via U-Net Based Architecture and Distributed Data Training, <https://www.sciopen.com/article/10.32604/cmcs.2025.060917>
52. M. B. Vibha, M. Chandrika, Samitha Khaiyum & P. Rakshitha Kiran, Predicting sedentary behavior in adults using stacked LSTM modeling, published: 27 November 2024
53. Volume 16, pages 346–355 (2025), <https://link.springer.com/article/10.1007/s13198-024-02622-2>
54. Bellary Chiterki Anil, Jayasimha Sondekoppa Rajkumar, T. L. Divya, Samitha khaiyum Rakshitha Kiran P., Balakrishnan Ramadoss, A Radiomics-based Framework for Liver Cancer Analysis using Explainable Artificial Intelligence (XAI) Methods, Vol. 15 No. 3 (2025): June, 2025, <https://etasr.com/index.php/ETASR/article/view/10377>
55. Jayminkumar Patel, Sanjaykumar patel, Birju Patel, Rajan shah, Development and Validation of a Stability-Indicating Analytical Method for Quantitative Determination of Selected Anticancer Drugs in Pharmaceutical Dosage Form, Archives / Vol. 35 No. S2 (2026): Apr-jun 2026, <https://welltestingjournal.com/index.php/WT/article/view/288>
56. Jayminkumar Patel, Sanjaykumar patel, Birju Patel, Rajan shah, Artificial Intelligence-Guided Optimization of Chromatographic Methods for Pharmaceutical Analytical Development, Archives / Vol. 35 No. S2 (2026): Apr-jun 2026, <https://welltestingjournal.com/index.php/WT/article/view/294>
57. Sharoon Sharif, Inikpi Egbunu, Yujian Fu, and Zhigang Xiao. 2026. Hybrid Multi-Modal Learning for Extreme Long-Tail Fungi Classification. In 2026 ACM Southeast Conference (ACMSE 2026), April 23–25, 2026, Troy, AL, USA. ACM, New York, NY, USA, 38–45. <https://doi.org/10.1145/3746467.3801502>