Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 07, NCIRTAI - 2026

# Blockchain and NFT-Based Solution for Genomic Data Management, Sharing, and Monetization

[S Kaviyapriya][1], [M Kaviya][2], [G Logeshwari][3], [R Dharanisri][4], [Gayatri], M.E.[5]
UG Scholar, Dept. of Computer Science and Engineering,
Jayalakshmi Institute of Technology, Tamilnadu, India[1234]
Assistant Professor, Dept. of Computer Science and Engineering,
Jayalakshmi Institute of Technology, Tamilnadu, India[5]

**Abstract**—Genomics has emerged as a transformative scientific field driven by advances in Next-Generation Sequencing (NGS), enabling rapid, accurate, and cost-effective analysis of complete genomes for personalized medicine and biomedical research. Existing genomic data management systems rely heavily on centralized repositories, which suffer from privacy risks, limited user control, lack of transparency, and absence of fair monetization mechanisms. This paper proposes GeneNFT, a decentralized blockchain- and NFT-based framework that introduces composable NFTs for genomic data ownership, in which a Raw Genomic Data (RGD) NFT acts as a parent asset for one or more Sequenced Genomic Data (SGD) NFTs, ensuring permanent traceability and data lineage. Three smart contracts automate ownership enforcement, consent-based data sharing, and transparent monetization. The system integrates the InterPlanetary File System (IPFS) for off-chain decentralized storage, Threshold Cryptography for distributed key management, and Fully Homomorphic Encryption (FHE) for privacy-preserving computation on encrypted genomic data. Experimental evaluation demonstrates 92% system accuracy and 95% security strength with practical execution latencies, confirming the framework's suitability for real-world deployment.

***Keywords*** - Genomic Data, NFT, Blockchain, Composable NFT, Threshold Cryptography, Fully Homomorphic Encryption, IPFS, Smart Contracts, Privacy-Preserving Computation, Decentralized Storage.

## I. INTRODUCTION

Genomics is the study of the complete set of genes (the genome) of organisms, encompassing their structure, function, evolution, and interactions. Next-Generation Sequencing (NGS) has become a powerful tool to identify genetic variants and variable gene expression patterns that correlate with disease states, enabling clinically relevant insights across oncology, rare diseases, infectious diseases, and prenatal diagnostics [1].

Genomic data sharing accelerates discoveries in personalized medicine and therapeutic development. However, existing systems are predominantly centralized, exposing data to single points of failure, data breaches, and lack of transparency. Individuals who contribute genomic data lose ownership and visibility into how it is accessed or monetized. Current systems also lack privacy-preserving mechanisms allowing secure analysis without revealing raw genetic information, discouraging participation and limiting dataset diversity.

This paper presents GeneNFT, a decentralized framework combining composable NFTs, smart contracts, IPFS-based decentralized storage, Threshold Cryptography, and Fully Homomorphic Encryption (FHE) to address these challenges. The system enables individuals to retain full ownership of their genomic data while allowing controlled sharing, secure computation, and transparent monetization.

The remainder of this paper is organized as follows: Section II reviews related work; Section III states the problem; Section IV describes the proposed system; Section V presents the system architecture; Section VI details technologies and algorithms; Section VII describes system modules; Section VIII outlines the working methodology; Section IX presents results; Section X discusses testing; and Section XI concludes.

## II. RELATED WORK

The intersection of blockchain, NFTs, and genomic data management has attracted significant research attention. Several studies have investigated how decentralized technologies can address privacy, ownership, and monetization challenges in genomic ecosystems.

### *A. Blockchain and NFT-Based Genomic Data Management*

Musamih et al. [1] proposed a blockchain and NFT-based solution for genomic data management on IEEE Access, demonstrating how NFTs can represent genomic dataset ownership while smart contracts automate access control and payment. GeneNFT extends this through composable NFT hierarchies and FHE integration.

### *B. NFT Frameworks for Digital Asset Ownership*

Srinivas and Pradhan [2] introduced TokenPharma, a novel NFT framework for pharmaceutical asset ownership and trading, demonstrating NFT-based ownership mechanisms beyond digital art. Jadon et al. [3] proposed an NFT-enhanced blockchain-based social network, confirming composable NFT structures can manage complex ownership hierarchies across diverse domains.

### *C. Blockchain-Based Genomic Databases*

Kimura et al. [4] assessed a blockchain-based genomic database for the Amazon Biobank, providing empirical evidence that blockchain storage is feasible for genomic data at scale, though without privacy-preserving computation or composable NFT ownership.

### *D. Privacy-Preserving Computation*

Baker et al. [10] investigated privacy-preserving linkage of genomic and clinical datasets. Gentry's foundational FHE work [15] provides the theoretical basis for the FHE module in GeneNFT,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 07, NCIRTAI - 2026

enabling direct computation on encrypted genomic data without decryption.

### *E. Threshold Cryptography*

Shamir's secret sharing scheme [14] underpins the Threshold Cryptography module in GeneNFT. By splitting private keys into distributed shares with a configurable reconstruction threshold, the system prevents any single entity from independently accessing genomic data, ensuring decentralized security and fault tolerance.

## III. PROBLEM STATEMENT

Existing genomic data systems are primarily centralized, storing sensitive genetic information in repositories controlled by third-party institutions. This model introduces single points of failure, vulnerability to data breaches, unauthorized access, and lack of transparency. Data owners have limited visibility into how their information is accessed or shared, raising serious privacy concerns.

Furthermore, the absence of transparent monetization models means individuals contributing genomic data receive no compensation despite the high commercial and research value of their data. The inability to compute on encrypted data further exacerbates privacy risks, as sensitive information must often be exposed during analysis. A decentralized, privacy-preserving system enabling controlled sharing, secure computation, and fair monetization is therefore urgently needed.

## IV. PROPOSED SYSTEM

GeneNFT is a decentralized web-based platform leveraging blockchain technology, composable NFTs, and advanced cryptographic techniques to provide complete genomic data ownership, controlled sharing, and transparent monetization. The system is designed around three core principles.

Privacy is achieved through Fully Homomorphic Encryption (FHE), enabling computation on encrypted genomic data without exposing raw genetic information. Threshold Cryptography ensures private encryption keys are distributed across independent nodes, preventing single-point compromise.

Ownership is enforced through a composable NFT hierarchy. A Raw Genomic Data (RGD) NFT represents the original genome upload; Sequenced Genomic Data (SGD) NFTs represent derived analysis outputs and are permanently linked to their parent RGD NFT, ensuring immutable data lineage.

Incentivization is realized through three smart contracts automating ownership verification, consent-based access control, and payment processing. Compensation is distributed directly to the data owner's blockchain wallet, creating a trustless economic model for genomic data sharing.

## V. SYSTEM ARCHITECTURE

GeneNFT follows a three-tier architecture comprising a Presentation Layer, an Application Layer, and an AI and Data Layer, ensuring modularity, maintainability, and scalability.

The Presentation Layer is built using HTML5, CSS3, Bootstrap 4, and JavaScript. It renders dashboards for data owners and researchers, including genome upload forms, dataset browsing, analysis request submission, and results visualization.

The Application Layer is implemented using Python with Flask. It handles HTTP routing, session management, user authentication, and coordinates communication between cryptographic modules, IPFS storage, the blockchain network, and the MySQL database.

The AI and Data Layer comprises the MySQL database, IPFS decentralized storage, blockchain network, and cryptographic algorithm pipeline. Raw genomic data is encrypted before upload to IPFS; the resulting CID is stored on-chain as NFT metadata. FHE-based analysis produces encrypted results that only authorized users can decrypt.

## VI. TECHNOLOGIES AND ALGORITHMS

### *A. Technology Stack*

The system uses Python 3.7.4, Flask 1.1.1, HTML5, CSS3, Bootstrap 4, JavaScript, and MySQL 5 on WampServer. Blockchain interaction uses Web3.py with MetaMask. Cryptographic operations use Pyfhel for FHE, shamir-mnemonic for threshold secret sharing, and PyCryptodome for AES-EAX encryption.

### *B. Composable NFT Architecture*

An RGD NFT is minted upon initial genome upload, encoding the IPFS CID, ownership details, timestamps, and access permissions. SGD NFTs minted for analysis outputs are linked to the parent RGD NFT via composability references in the smart contract, ensuring permanent traceability and immutable data lineage across all derived datasets.

### *C. Threshold Cryptography*

A 3-of-5 Shamir secret sharing scheme splits the private key into five shares, any three of which suffice to reconstruct the key. Each share is encrypted with a node-specific Fernet key before storage, preventing any single node from independently accessing genomic data while ensuring fault tolerance against node compromise.

### *D. Fully Homomorphic Encryption (FHE)*

FHE allows genomic analysis -- disease risk prediction, mutation classification, variant identification -- to be performed directly on encrypted data using the Pyfhel library (BFV/CKKS scheme). All intermediate computations remain encrypted; only the final result is decrypted by the authorized requester.

### *E. IPFS Content Addressing*

IPFS stores large genomic files off-chain using content-based addressing. Each encrypted file receives a unique CID computed from its hash, ensuring data integrity. The CID is stored on-chain as NFT metadata, enabling tamper-proof referencing without on-chain storage overhead.

### *F. Smart Contract Execution*

Three smart contracts govern system operations: the Ownership Contract verifies NFT ownership; the Access Control Contract manages consent-based per-NFT permissions; the Monetization Contract validates payment and distributes compensation to the owner's wallet. All contracts execute autonomously without intermediaries.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 07, NCIRTAI - 2026

## VII. SYSTEM MODULES

### *A. User Registration and Authentication*

During registration, an RSA public/private key pair is generated, the private key is split into 3-of-5 threshold shares distributed across nodes, and the public key and share hashes are stored on-chain. Each user receives a unique blockchain wallet address for tokenized transactions.

### *B. Key Management and Threshold Reconstruction*

Private keys are split into Shamir shares, each encrypted with node-specific Fernet keys and stored in the database. When authorized access is requested, the module collects required shares, verifies the threshold condition via smart contract, and temporarily reconstructs the private key for controlled decryption.

### *C. Genome Upload and Encryption*

Data owners upload genomic files after public key verification. Hybrid encryption is applied: the file is AES-EAX encrypted with a randomly generated symmetric key, which is itself RSA-encrypted. The encrypted file is uploaded to IPFS, generating a CID anchored to the blockchain via NFT minting.

### *D. NFT Minting and Composability*

An RGD NFT is minted with metadata including IPFS CID, owner wallet address, dataset title, permitted analysis types, pricing, and timestamp. An SGD NFT is minted for each analysis output and linked to its parent RGD NFT, preserving data lineage and enabling hierarchical ownership tracking.

### *E. Dataset Discovery and Analysis Request*

Researchers search available datasets by disease type. The module queries VCF files for matching genetic variants and returns metadata including pricing and non-sensitive owner identity. Researchers submit analysis requests specifying dataset ID, analysis type, and research purpose; the smart contract layer validates ownership, consent, and payment.

### *F. Privacy-Preserving Computation and Result Delivery*

Authorized analysis is performed on encrypted genomic data using FHE, producing encrypted outputs such as disease risk scores, mutation lists, and VCF files. An SGD NFT is minted for each output. Results are delivered via the web interface with access enforced by smart contract.

### *G. Audit and Logging*

All system events -- registrations, uploads, NFT minting, access requests, approvals, computations, and transactions -- are recorded as immutable blockchain logs, ensuring transparency, accountability, and regulatory compliance.

## VIII. WORKING METHODOLOGY

The GeneNFT operational workflow proceeds through nine steps: (1) Requirement setup; (2) MySQL database design for users, datasets, NFT records, key shares, transactions, and audit logs; (3) Frontend development with owner and researcher dashboards; (4) Flask backend APIs for authentication, upload, encryption, IPFS, blockchain, and FHE; (5) Integration of Threshold Cryptography, FHE, IPFS, NFT minting, and smart contract modules; (6) Unified system integration via APIs; (7) Smart contract deployment via Web3.py; (8) Unit, integration, functional, security, performance, and usability testing; (9) Deployment with IPFS nodes and blockchain network configured for real-time operation.

## IX. EXPERIMENTAL RESULTS

GeneNFT is evaluated across six performance metrics measuring efficiency, security, and accuracy on Windows 10 with VS Code, WampServer, Web3 with MetaMask, and IPFS.

**TABLE I. SYSTEM PERFORMANCE METRICS**

| Metric | Value |
|---|---|
| Data Encryption Time | 2.5 sec |
| Key Reconstruction Time | 1.8 sec |
| Smart Contract Execution | 3.2 sec |
| FHE Computation Time | 4.5 sec |
| System Accuracy | 92% |
| Security Strength | 95% |

Data encryption time averages 2.5 seconds, indicating efficient handling of large genomic datasets. Key reconstruction time of 1.8 seconds confirms that threshold cryptography provides secure access with minimal delay. Smart contract execution time of 3.2 seconds reflects acceptable blockchain latency for validation and payment processing.

FHE-based computation on encrypted data requires 4.5 seconds due to the additional overhead of privacy-preserving operations. System accuracy of 92% confirms reliable analytical results on encrypted data. Security strength of 95% validates multi-layered protection through encryption, threshold key management, and blockchain access control.

## X. SYSTEM TESTING

### *A. Test Coverage*

Five testing categories are conducted: (1) Functional -- user registration, login, genome upload, analysis request, and NFT minting all pass; (2) Security -- invalid logins rejected, unauthorized access blocked, encrypted data protected, and insufficient threshold key shares denied; (3) Performance -- acceptable execution times for upload, encryption, smart contract execution, and FHE computation; (4) Integration -- correct data flow between IPFS, blockchain, frontend, and backend; (5) Usability -- smooth navigation, form submission, and results visualization.

### *B. Bug Report*

**TABLE II. BUG REPORT SUMMARY**

| BID | TC | Description | Status | Result |
|---|---|---|---|---|
| B01 | TC02 | Session handling on login | Fixed | Pass |
| B02 | TC11 | Encryption delay large file | Fixed | Pass |
| B03 | TC18 | Contract not triggering | Fixed | Pass |

All 25 test cases passed after bug resolution. The system demonstrates reliable handling of sensitive genomic data with proper security mechanisms and is ready for real-world deployment.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 07, NCIRTAI - 2026

## XI. CONCLUSION AND FUTURE WORK

This paper presented GeneNFT, a comprehensive decentralized framework for genomic data management integrating composable NFTs, Threshold Cryptography, Fully Homomorphic Encryption, IPFS-based storage, and automated smart contracts. The system effectively overcomes centralized limitations by ensuring data ownership, controlled access, transparency, and fair monetization without trusted third parties.

The RGD-SGD composable NFT hierarchy provides permanent traceability and immutable data lineage. The 3-of-5 Threshold Cryptography scheme prevents single-point key compromise. FHE enables genomic analysis without ever exposing raw genetic information. Experimental evaluation confirms 92% system accuracy, 95% security strength, and practical execution latencies.

Future work will integrate with Electronic Health Record (EHR) systems for combined genomic-clinical analysis, incorporate advanced AI and deep learning models for disease prediction and mutation classification, and explore zero-knowledge proofs for additional privacy guarantees. Pilot deployments with healthcare institutions will measure real-world impact on genomic data governance and research outcomes.

## REFERENCES

[1] A. Musamih et al., "Blockchain and NFT-based solution for genomic data management, sharing, and monetization," IEEE Access, vol. 13, pp. 35780-35810, 2025.
[2] O. Srinivas and N. R. Pradhan, "A novel NFT framework for pharmaceutical asset ownership: TokenPharma," IEEE Access, vol. 12, pp. 171418-171430, 2024.
[3] S. Jadon et al., "Non-fungible token enhanced blockchain-based online social network," IEEE Access, vol. 12, pp. 92368-92380, 2024.
[4] L. T. Kimura et al., "Amazon biobank: Blockchain-based genomic database," IEEE Access, vol. 12, pp. 9632-9645, 2024.
[5] G. C. Lee, H. Y. Koo, and H. Lee, "Quadratic regression models for NFT valuation," IEEE Access, vol. 13, pp. 114029-114040, 2025.
[6] A. Battah et al., "Blockchain and NFTs for trusted ownership of AI models," IEEE Access, vol. 10, pp. 112230-112250, 2022.
[7] A. C. Moreaux and M. P. Mitrea, "Royalty-friendly digital asset exchanges on blockchains," IEEE Access, vol. 11, pp. 56235-56250, 2023.
[8] S. A. Gebreab et al., "NFT-based traceability for medical devices," IEEE Access, vol. 10, pp. 120000-120015, 2022.
[9] X. Liu et al., "Federated learning with blockchain for genomic analysis," IEEE Access, 2021.
[10] D. B. Baker et al., "Privacy-preserving linkage of genomic and clinical datasets," IEEE Trans. Comput. Biol., vol. 16, no. 5, pp. 1500-1510, 2019.
[11] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
[12] V. Buterin, "Ethereum white paper," 2014.
[13] J. Benet, "IPFS -- Content addressed distributed file system," 2014.
[14] A. Shamir, "How to share a secret," Commun. ACM, 1979.
[15] C. Gentry, "Fully homomorphic encryption using ideal lattices," in Proc. STOC, 2009.
[16] J. Shendure and H. Ji, "Next-generation DNA sequencing," Nature Biotechnology, 2008.
[17] HIPAA, "Health Insurance Portability and Accountability Act," 1996.
[18] GDPR, "General Data Protection Regulation," European Union, 2018.