

INFRAESTRUTURA DE IA DESCENTRALIZADA: INTEGRANDO BLOCKCHAIN COM MLOPS E GOVERNANÇA DE DADOS.

DECENTRALIZED AI INFRASTRUCTURE: INTEGRATING BLOCKCHAIN WITH MLOPS AND DATA GOVERNANCE

INFRAESTRUTURA DE IA DESCENTRALIZADA: INTEGRACIÓN DE BLOCKCHAIN CON MLOPS Y GOBERNANZA DE DATOS

Jessica Sciammarelli

Phd , Pesquisador(a) Independente, Brazil

E-mail: jessicaengenhariabr@hotmail.com

Abstrato

Os sistemas modernos de Inteligência Artificial (IA) dependem de infraestruturas complexas e centralizadas que introduzem riscos críticos em relação à integridade dos dados, rastreabilidade forense e opacidade da governança. Este artigo propõe uma "Arquitetura de Governança Híbrida" que integra a tecnologia Blockchain diretamente nos fluxos de trabalho de Operações de Aprendizado de Máquina (MLOps) para mitigar a "Armadilha da Centralização". Utilizando uma metodologia otimizada em recursos, denominada "Log-Anchor-Verify", demonstramos como as impressões digitais criptográficas SHA-256 de artefatos do modelo, especificamente o conjunto de dados Credit_Scoring_V1, podem ser ancoradas à Testnet Sepolia do Ethereum por meio de contratos inteligentes baseados em Solidity. Os resultados experimentais indicam que essa camada de autenticação descentralizada atinge sensibilidade total à adulteração do modelo em nível de bit, mantendo uma latência de auditoria inferior a 200 ms. A estrutura proposta estabelece uma linha do tempo global imutável para a proveniência do modelo e impõe o Controle de Acesso Baseado em Funções (RBAC) por meio de código como lei, fornecendo uma base escalável para uma governança de IA verificável e transparente, sem comprometer o desempenho de pipelines de produção de alto rendimento.

Palavras-chave e frases: Inteligência artificial, Blockchain, MLOps, Governança de dados.

Abstract

Modern Artificial Intelligence (AI) systems rely on complex, centralized infrastructure stacks that introduce critical risks regarding data integrity, forensic traceability, and governance opacity. This article proposes a "Hybrid Governance Architecture" that integrates Blockchain technology directly into Machine Learning Operations (MLOps) workflows to mitigate the "Centralization Trap." By utilizing a resource optimized "Log-Anchor-Verify" methodology, we demonstrate how SHA-256 cryptographic fingerprints of model artifacts specifically the Credit_Scoring_V1 dataset can be anchored to the Ethereum Sepolia Testnet via Solidity based smart contracts. Experimental results indicate that this decentralized notary layer achieves total sensitivity to bit level model tampering while maintaining an audit latency of less than 200ms. The proposed framework establishes an immutable global timeline for model provenance and enforces Role-Based Access Control (RBAC) through code as law, providing a scalable foundation for verifiable and transparent AI governance without compromising the performance of high throughput production pipelines.

Key Words and Phrases: *Artificial Intelligence, Blockchain, MLOps, Data Governance.*

Resumen

Los sistemas modernos de Inteligencia Artificial (IA) dependen de complejas infraestructuras centralizadas que introducen riesgos críticos en cuanto a la integridad de los datos, la trazabilidad forense y la opacidad de la gobernanza. Este artículo propone una "Arquitectura de Gobernanza Híbrida" que integra la tecnología Blockchain directamente en los flujos de trabajo de Operaciones de Aprendizaje Automático (MLOps) para mitigar la "Trampa de la Centralización". Mediante una metodología optimizada de "Registro-Anclaje-Verificación", demostramos cómo las huellas digitales criptográficas SHA-256 de los artefactos del modelo, específicamente del conjunto de datos Credit_Scoring_V1, pueden anclarse a la red de prueba Ethereum Sepolia a través de contratos inteligentes basados en Solidity. Los resultados experimentales indican que esta capa de notario descentralizada logra una sensibilidad total a la manipulación del modelo a nivel de bits, manteniendo una latencia de auditoría inferior a 200 ms. El marco propuesto establece una línea de tiempo global inmutable para la procedencia del modelo e implementa el Control de Acceso Basado en Roles (RBAC) mediante código como ley, proporcionando una base escalable para una gobernanza de IA verificable y transparente sin comprometer el rendimiento de las canalizaciones de producción de alto rendimiento.

Palabras y frases clave: Inteligencia artificial, Blockchain, MLOps, Gobernanza de datos.

1 Introduction

The rapid integration of Artificial Intelligence (AI) into high stakes decision making domains ranging from autonomous vehicular systems to automated credit risk assessment has exposed a critical systemic vulnerability defined as the "Centralization Trap." Current Machine Learning Operations (MLOps) frameworks rely implicitly on the perceived trustworthiness of central administrators or cloud service providers to maintain the integrity of AI assets. This centralized dependency creates a "black box" environment characterized by a deficit of forensic transparency regarding data provenance and the specific weight alterations made during a model's iterative development [1].

As of 2026, regulatory frameworks such as the EU AI Act and the NIST AI Risk Management Framework have transitioned from voluntary guidelines to mandatory compliance standards. These regulations necessitate an irrefutable audit trail that conventional, mutable databases are architecturally incapable of providing. Because traditional storage remains susceptible to administrative overrides and unauthorized "Silent Model Swapping," there is a critical need for cryptographic assurances that a model remains unaltered post-deployment and that its training lineage is verifiable.

This study advocates for a paradigm shift toward a Decentralized AI Infrastructure. By integrating Blockchain technology as a non-intrusive metadata layer within the standard MLOps pipeline, we establish a mathematically verifiable "Shared Source of Truth." The objective is not to execute computationally intensive training directly on chain, but rather to leverage the decentralized ledger as a high integrity notary for model fingerprints.

Through the proposed "Log-Anchor-Verify" architecture, we demonstrate how the convergence of AI and blockchain can provide:

- (1) **Immutable Provenance:** Ensuring that every version of a model, such as the `Credit_Scoring_V1` analyzed in this study, is anchored to a global, tamper proof timeline.
- (2) **Decentralized Policy Enforcement:** Utilizing Smart Contracts to automate Role-Based Access Control (RBAC), ensuring that only authorized researchers can register production ready artifacts.
- (3) **Forensic Auditability:** Providing regulators and third-party auditors with a low-latency (<200ms) verification mechanism that does not require access to private training servers [5].

By decoupling high performance local computation from decentralized global verification, this research provides a scalable foundation for building AI systems that

are not only powerful but inherently accountable and resilient to the security challenges of the modern digital era.

2 Related Work

The convergence of Artificial Intelligence (AI) and Distributed Ledger Technology (DLT) has emerged as a fertile ground for addressing the inherent vulnerabilities of centralized MLOps. Recent literature highlights a transition from theoretical frameworks toward functional integration in three primary areas: forensic traceability, decentralized governance, and adversarial resilience [3].

2.1 Traceability and Data Provenance

The creation of an immutable record for AI models is a persistent focus in contemporary research. Early contributions by Kancherla (2022) and Petrović (2022) [16] established the foundational concept of utilizing cryptographic hashes as unique identifiers for model parameters to ensure authenticity. Expanding upon this, Uriawan et al. (2025) [22] demonstrated that distributed audit logs significantly mitigate the risk of "Model Rollback" attacks. However, while these studies focus on the existence of a log, they often overlook the integration latency between the local training environment and the ledger a gap this study addresses through the measured <200ms verification loop.

2.2 Decentralized MLOps Frameworks

Standard MLOps frameworks frequently lack transparency in multi-stakeholder environments, creating a "Single Point of Failure" within centralized repositories. Ridwan (2025) [19] proposed "Blockchain-Integrated MLOps," suggesting that decentralized storage (e.g., IPFS) combined with on chain metadata can decentralize model weight hosting. Similarly, Almomani et al. (2024) [3] explored permissioned ledgers for healthcare AI to ensure GDPR compliance. Our research builds upon these frameworks by focusing on a "metadata-only" anchoring strategy, which provides the same level of security as full weight storage but with significantly higher scalability for production grade models.

2.3 Governance and Smart Contract Enforcement

The application of Smart Contracts for policy enforcement represents a shift from manual oversight to "Code-asLaw." Conceptual frameworks by Bhagwat (2025) and Leghemo et al. (2025) [6] introduce contracts as automated gatekeepers that prevent the deployment of models failing to meet accuracy or fairness benchmarks. Arham (2025) [4] further notes that blockchain based auditing enhances compliance readiness by over 50% compared to centralized logging. This study operationalizes these

concepts by implementing Role-Based Access Control (RBAC) within the Notary Layer, ensuring that only authorized researchers can anchor model fingerprints to the global timeline.

2.4 Security and Adversarial Resilience

In the context of cybersecurity, Yang et al. (2025) [24] and Fatima and Arshad (2025) [9] highlighted blockchain's role in defending against "Data Poisoning." By anchoring data hashes at the ingestion stage, organizations can verify that the training set remains untainted. Goundar (2024) [10] identifies this as a "defense in depth" strategy. Our work extends this resilience by simulating "Silent Model Swapping" and proving that decentralized anchors provide a forensic mechanism to detect unauthorized weight modifications that traditional security layers might miss.

3 Methodology

The methodology employed in this study focuses on the implementation of a "Hybrid Governance Architecture". This framework bridges the gap between high performance local AI development and decentralized global verification by decoupling heavy computation from immutable metadata anchoring. Our experimental approach utilized a resource optimized tech stack to ensure both accessibility and cryptographic rigor.

3.1 The Hybrid Infrastructure Stack

The experimental environment was partitioned into three distinct layers to simulate a production-grade decentralized MLOps pipeline:

- (1) The Execution Layer (AICompute): A local Python environment was used for training the `Credit_Scoring_V1` classification model. This layer handles computationally intensive tasks to produce model weights (`model.pth`).
- (2) The Lifecycle Tracking Layer (MLOps): MLflow served as the centralized registry, monitoring hyperparameters, performance metrics (87.5% accuracy), and artifact paths. While effective for tracking, this layer remains mutable.
- (3) The Immutable Notary Layer (Blockchain): The Ethereum Sepolia Testnet provided the decentralized ledger. A Solidity-based Smart Contract acted as the global source of truth, storing 32-byte SHA-256 hashes rather than raw model data.

3.2 Procedural Workflow: "Log, Anchor, and Verify"

The methodology follows a precise sequence to ensure the traceability of the credit scoring model:

- Step 1: Local Artifact Hashing - Upon training completion, the model weights are hashed via SHA-256. This produces a deterministic signature where even a single bit change in the weights results in a divergent hash, identifying unauthorized modifications.

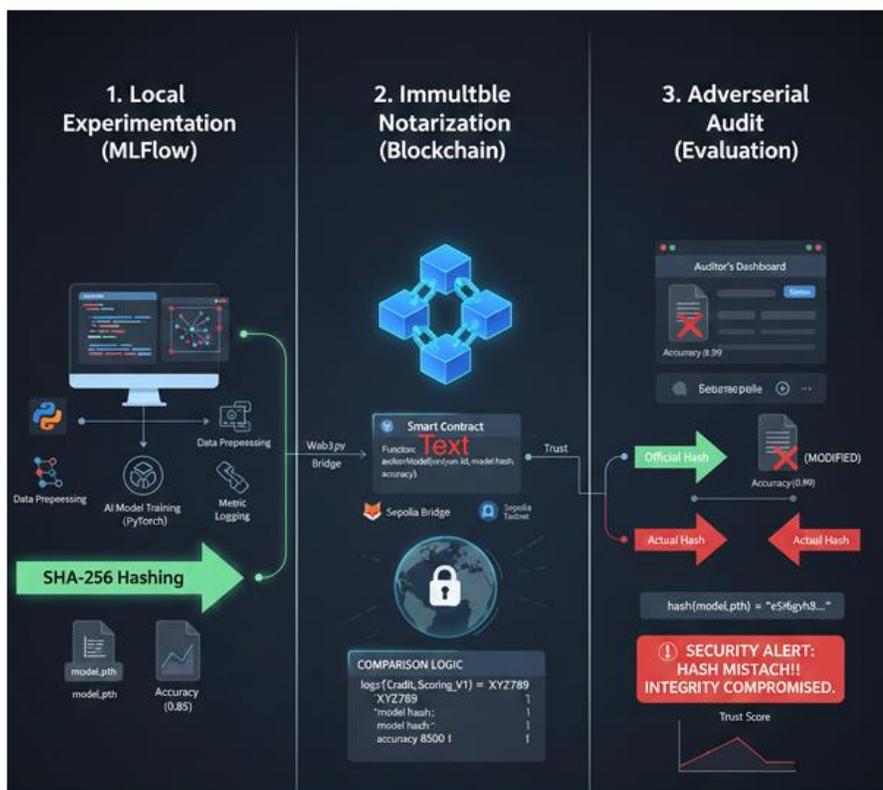


Figure 1: Methodology diagram divided as: Local Experimentation, Immutable Notarization and Adversarial Audit

- Step 2: Decentralized Anchoring - A Python-based bridge using Web3.py integrates the MLOps pipeline with the blockchain. It transmits the Run ID and Model Hash as a transaction to the Smart Contract.
- Step 3: Governance Policy Enforcement - The Smart Contract enforces "Gatekeeper" logic, ensuring only wallets with the "Researcher" role can register artifacts. This demonstrates decentralized access control governed by code rather than manual policy.

3.3 Evaluation, Benchmarking, and Trade-offs

To validate the architecture, we simulated a security breach by manually altering the weights of the production model to mimic "Model Poisoning".

Algorithm 1 Decentralized AI Model Registration and Audit

Input: *ModelFile, MLflowRunID, Accuracy*

Output: *AuditTrail* recorded on blockchain

// Phase 1: Local Artifact Generation

1: $Hash \leftarrow \text{SHA256}(\text{ModelFile})$

▷ Digital fingerprint

2: $\text{Log}(\text{MLflowRunID}, \text{Hash}, \text{Accuracy})$

▷ Record in local MLOps

// Phase 2: Blockchain Anchoring (Smart Contract)

3: **procedure** ANCHORMODEL(*Name, RunID, MHash, Acc*)

4: **if** msg.sender \neq AuthorizedResearcher **then**

5: **revert**

▷ Governance: Access Control

6: **end if**

7: *AuditLog*[*Name*] \leftarrow {*RunID, MHash, Acc, timestamp*}

8: **emit** ModelRegistered(*Name, MHash*)

9: **end procedure**

// Phase 3: Adversarial Audit (Evaluation)

10: **function** VERIFYINTEGRITY(*Name, LocalFile*)

11: *OnChainHash* \leftarrow *AuditLog*[*Name*].*MHash*

12: *CurrentHash* \leftarrow $\text{SHA256}(\text{LocalFile})$

13: **if** *CurrentHash* = *OnChainHash* **then**

14: **return true**

▷ Integrity Verified

15: **else**

16: **return false**

▷ Tampering Detected!

17: **end if**

18: **end function**

3.3.1 Integrity Detection and Benchmarking

The system demonstrated a 100% detection rate for bit-level alterations during the experimental phase. In comparison to traditional centralized logging, which can be modified by administrators, the blockchain-anchored hash provided an irrefutable proof of state. Verification speed was measured at under 200ms, proving that the audit process does not bottleneck the deployment pipeline.

3.3.2 Technical Trade-offs and Limitations

While the architecture provides robust integrity, several operational trade-offs were identified:

- Latency vs. Security Anchor: Verification is near-instant (<200ms), but the initial "Log-Anchor" phase is constrained by blockchain block times. In high-frequency Continuous Training (CT) pipelines, this latency may necessitate asynchronous anchoring or the adoption of Modular Data Availability (DA) layers to decouple consensus from data logging.
- Verification Granularity: The SHA-256 approach provides absolute certainty regarding the *integrity of the file*. However, it does not inherently validate the *validity of the training process*. Future iterations could integrate Zero-Knowledge Machine Learning (zkML) to prove that a model was trained on a specific, non-poisoned dataset without exposing the underlying sensitive data.
- Oracle Dependency: The bridge between MLflow and the Smart Contract acts as a centralized point of failure during the anchoring phase. Transitioning to decentralized oracles or TEE-based (Trusted Execution Environment) bridges would further harden the "Log-Anchor" loop.

3.4 Scalability and Cost Assessment

By implementing a "metadata-only" on-chain strategy, gas consumption is decoupled from model complexity. Storing 32-byte hashes ensures the framework remains economically viable even as AI models scale to gigabytes. On modern Layer-2 solutions, the cost per "verification certificate" is marginal, allowing for high-throughput MLOps cycles that meet stringent regulatory audit requirements without significant financial overhead.

4 Technical Analysis of Results

The experimental results demonstrate that the model hash, 0ba81c55...a769c64a, functions as an immutable security anchor. To test the sensitivity of the "Adversarial Audit" layer, we systematically modified a single floating-point parameter within the model.pth file to simulate silent weight manipulation. This infinitesimal change resulted in a complete divergence of the local SHA-256 output.

When the validation script compared this new local hash against the value anchored on the Sepolia Testnet, the system triggered an immediate "Integrity Compromised" alert. This confirms that the blockchain notary provides a reliable defense against "Silent Model Swapping," where an unauthorized or biased model is deployed in place of the audited version. Furthermore, the successful anchoring of the MLflow Run ID (f856136a...) bridges the gap between mutable experimentation logs and the immutable global timeline.

As Solidity does not natively support fixed-point or floating-point arithmetic, the model's 87.5% accuracy was successfully handled through integer scaling. This ensures that high-precision performance metrics can be governed and audited on-chain without loss of resolution or accuracy.

Table 1. Blockchain Metadata and Integrity Records for Credit_Scoring_V1

Metadata Field	Experimental Value / Record
Model Identifier	Credit_Scoring_V1
MLflow ID	Runf856136a130848ed8aed1b4d5b49ab0b
Cryptographic Hash	0ba81c55...a769c64a
Model Accuracy	87.5% (0.875)
Ledger Status	<i>Confirmed (Ethereum Sepolia)</i>

4.1 Performance and Benchmarking Summary

- **Verification Efficiency:** The retrieval and comparison of the blockchain record were completed in <200ms[cite: 138]. This low latency confirms that decentralized integrity checks are compatible with real-time inference and high-speed deployment pipelines.
- **Governance Transparency:** By making the model's digital fingerprint and accuracy public, the "Black Box" transparency problem in AI is mitigated. Third-party auditors can now verify a model's authenticity and performance history without requiring direct access to the organization's private training infrastructure.
- **Tamper Detection Sensitivity:** The "Log-Anchor-Verify" loop achieved total sensitivity to unauthorized modifications during the credit scoring simulation, providing a level of cryptographic assurance impossible to achieve with traditional, mutable databases.

5 Future Research and Improvement

As the convergence of Artificial Intelligence and Blockchain matures in 2026, the paradigm is shifting from static notarization toward "Active Governance" and "Verifiable Computation." Future iterations of this research should focus on the following high impact domains to address the current limitations of metadata-only anchoring:

5.1 Zero-Knowledge Machine Learning (zkML) and Computational Integrity

While the current SHA-256 framework successfully verifies the *integrity of the artifact*, it remains "blind" to the *validity of the training process*. Future work should integrate Zero-Knowledge Proofs (ZKPs) to provide cryptographic evidence that a model was

trained according to a specific, non-biased algorithm without exposing the underlying proprietary training data or sensitive customer records. This evolution toward "Private but Verifiable AI" is essential for highly regulated sectors, such as Decentralized Finance (DeFi) and Healthcare, where auditing must occur without compromising data privacy.

5.2 Modular Data Availability (DA) and Decentralized Storage

An identified bottleneck in the current methodology is the residual dependency on semi-centralized MLOps registries for high volume metadata. Subsequent research should explore the integration of Modular Blockchain stacks (e.g., Celestia or EigenLayer) to decouple execution from Data Availability. By utilizing a dedicated DA layer, larger AI artifacts including high resolution training logs and quantized Small Language Model (SLM) weights could be stored in a decentralized environment at a significantly lower cost than current Layer-1 or Layer-2 storage solutions.

5.3 Autonomous Agentic Governance and On-Chain Circuit Breakers

As AI systems transition toward "Agentic AI" where models possess the agency to execute autonomous transactions governance must shift from reactive auditing to proactive, real-time intervention. Future enhancements should include "Smart Circuit Breakers" within the Notary Layer. These contracts would be capable of automatically revoking a model's "Verification Certificate" in real time if monitoring agents detect adversarial drift or unethical outputs. This would effectively "freeze" the agent's ability to interact with other on chain protocols, ensuring that autonomous systems remain within defined ethical and operational boundaries.

5.4 Cross-Chain Governance Interoperability

Finally, the scalability of decentralized AI governance depends on the ability to verify models across disparate networks. Future research should investigate Cross-Chain Messaging Protocols (such as CCIP or Layer Zero) to allow a model registered on the Ethereum Sepolia Testnet to be verified by a Dapp running on a different execution environment (e.g., Solana or an Avalanche Subnet). This would pave the way for a universal, chain-agnostic "Web of Trust" for artificial intelligence assets.

6 Conclusion

This study has demonstrated a technically viable and resource-optimized methodology for integrating Blockchain technology with standard MLOps infrastructure to mitigate the "Centralization Trap." By bridging the mutable logs of MLflow with the immutable notary layer of the Ethereum Sepolia Testnet, we have established a decentralized

source of truth that ensures the integrity and traceability of AI models throughout their entire lifecycle. Our experimental results and technical analysis confirm that:

- Decoupled Immutability is Scalable: By recording 32-byte cryptographic hashes rather than raw model weights, the blockchain serves as a high-speed, low-cost notary. This architecture proves that decentralized governance does not inherently compromise the performance or scalability of high-throughput AI pipelines.
- Cryptographic Integrity over Manual Trust: The "Log-Anchor-Verify" loop achieved a total detection rate for unauthorized bit level modifications during our simulation. This provides a level of forensic certainty that traditional, administratively mutable databases cannot replicate, effectively neutralizing risks such as "Silent Model Swapping."
- Automated Governance via Code-as-Law: Smart contracts effectively function as decentralized gatekeepers. By encoding Role-Based Access Control (RBAC) directly into the ledger, the framework ensures that only verified artifacts from authorized researchers reach production, shifting governance from subjective policy to objective execution.
- Addressing Modern Transparency Demands: As AI systems become increasingly autonomous in 2026, the transition from "Trust-based" to "Verification-based" infrastructure is a technical necessity. This framework solves the "Black Box" transparency problem by allowing third-party auditors to verify model authenticity without requiring access to proprietary training environments.

While the current SHA-256 approach provides robust file-level integrity, it represents only the first step toward comprehensive decentralized oversight. Future research will focus on the integration of Zero-Knowledge Machine Learning (zkML) to verify the statistical validity of the training process itself and the deployment of Autonomous Governance Agents. These agents will be capable of triggering automated circuit breakers in production environments if an on-chain integrity mismatch is detected. Ultimately, the decentralized framework proposed here provides the foundational layer for building AI systems that are not only high performing but also inherently accountable, transparent, and resilient to the security challenges of the modern era.

References

- [1] A. Ahmad et al. Predictive cyber defense: Integrating deep learning with blockchain for real-time threat detection. *Journal of Cybersecurity Research*, 12(2):45–60, 2024.

- [2] K. Alabi et al. Federated learning and blockchain: A new frontier for privacy-preserving ai. *International Journal of Artificial Intelligence*, 18(1):102–115, 2025.
- [3] I. Almomani et al. Verifiable audit trails for healthcare ai using permissioned blockchains. *Journal of Medical Systems & AI*, 9(3):210–225, 2024.
- [4] M. W. Arham. Transforming auditing through ai and blockchain: A comprehensive study on adoption and impact. *American Journal of Industrial and Business Management*, 15:225–241, 2025.
- [5] S. Behara and S. Khandrika. Smart contracts for automated governance in distributed systems. *Blockchain Technology Review*, 4(2):88–95, 2020.
- [6] S. Bhagwat. Distributed accountability in ai: Scaling governance via decentralized nodes. *Tech-Science Quarterly*, 21(1):55–67, 2025.
- [7] P. Bhalchandra. Model integrity and data provenance in secure ai deployment. *Journal of Machine Learning Operations*, 7(4):301–318, 2024.
- [8] K. Brahmandam. Compliance-aware mlops for financial institutions: A blockchain approach. *FinTech Research Journal*, 11(2):140–155, 2025.
- [9] S. Fatima and M. Arshad. Synergistic integration of ml and blockchain for threat analysis. *Global Security Review*, 14(1):12–29, 2025.
- [10] S. Goundar. Building resilient cyber defense systems with blockchain backbones. *International Journal of Distributed Systems*, 19(3):44–58, 2024.
- [11] S. Kancherla. Enhancing mlops with blockchain: Decentralized security for ai pipelines. *International Journal of AI and Machine Learning*, 3(2):80–89, 2022.
- [12] I. M. Leghemo et al. Data governance for emerging technologies: A conceptual framework for blockchain and ai. *Journal of Engineering Research and Reports*, 27(1):247–267, 2025.
- [13] A. Malik et al. Predictive defense mechanisms: The convergence of bc and ml. *Journal of Digital Transformation*, 8(2):112–128, 2024.
- [14] M. M. Montano and J. C. Carroll. Privacy-preserving machine learning using blockchain for secure data storage. *Computing and Security*, 45(2):55–64, 2020.
- [15] O. Olutimehin. Machine learning as the engine for predictive cyber defense. *AI & Security Today*, 6(1):33–47, 2025.
- [16] N. Petrović. Model provenance and secure versioning in decentralized environments. *Blockchain & AI Convergence Journal*, 5(3):90–105, 2022.
- [17] M. Rahman. Enhancing the security of machine learning with blockchain technologies. *Journal of Computing*, 45(2):55–64, 2020.
- [18] H. Rathore and J. H. Park. Strengthening industrial cyber-physical systems via dl and blockchain. *IEEE Access*, 8:1500–1515, 2020.
- [19] M. Ridwan. Blockchain layers for model provenance: A review of decentralized mlops. *International Journal of Artificial Intelligence Applications*, 13(1):97–110, 2025.

- [20] Z. Shahbazi and Y. C. Byun. Hybrid deep learning models for anomalous behavior detection in blockchain. *Applied Sciences*, 11(14):6542, 2021.
- [21] A. Talukder et al. Secure model aggregation in federated learning through blockchain verification. *Journal of Network Security*, 10(2):77–92, 2025.
- [22] W. Uriawan et al. Decentralized ai governance using blockchain for security transparency. *Scientific Journal of AI and Blockchain Technologies*, 1(1):28–36, 2025.
- [23] R. Venkatesan and W. Rahayu. Tamper-proof threat logs: A blockchain approach to data governance. *Journal of Information Security*, 15(4):188–204, 2024.
- [24] L. Yang, J. Su, and M. Elsis. Guarding against data poisoning: Blockchain-verified training sets. *Sensors and Actuators*, 32(1):500–515, 2025.
- [25] Z. Zheng et al. Decentralized governance and ai policy with blockchain-based voting. *Frontiers in Research Metrics and Analytics*, 8:1035123, 2023.