iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

BLOCKCHAIN-GOVERNED DATA LINEAGE AND PROVENANCE FOR FINANCIAL TRANSACTION MONITORING

Vikas Reddy Mandadhi
Bellevue University,
vmandadhi@my365.bellevue.edu

ABSTRACT

The increasing complexity of digital financial ecosystems has amplified the need for transparent, tamper-proof, and auditable data trails. Traditional financial transaction monitoring systems struggle to maintain consistent data lineage due to siloed databases, fragmented processing pipelines, and limited visibility across institutional boundaries. These limitations weaken compliance efforts in Anti-Money Laundering (AML), Counter-Terrorist Financing (CFT), and regulatory reporting. This paper proposes a blockchain-governed framework for capturing end-to-end data lineage and provenance in financial transaction monitoring. The model integrates metadata extraction, cryptographic hashing, permissioned distributed ledgers, and smart contract–driven governance to ensure immutable, time-stamped, and verifiable transaction histories. The framework provides a unified provenance layer that enhances traceability, reduces audit complexity, and improves regulatory oversight. A structured evaluation compares traditional monitoring systems with the proposed blockchain-based architecture, highlighting improvements in transparency, tamper resistance, and automated compliance reporting. The study contributes a practical governance model, a technical workflow for blockchain-enabled lineage, and insights into institutional adoption challenges. This research demonstrates that blockchain-governed data provenance can significantly strengthen financial risk management, fraud detection, and regulatory compliance in modern financial infrastructures.

Keywords:

Blockchain, Data Lineage, Data Provenance, Financial Transactions, AML, Compliance, Distributed Ledger Technology.

1. INTRODUCTION

1.1 Background

The modern financial ecosystem has evolved into a highly dynamic, interconnected, and data-intensive environment. Financial institutions process millions of transactions daily across digital banking platforms, cross-border payment networks, mobile money systems, and automated settlement infrastructures. This exponential increase in transaction volume and structural complexity poses significant challenges for maintaining transparency, auditability, and trust in financial processes.

Regulatory bodies worldwide—such as the Financial Action Task Force (FATF), Basel Committee on Banking Supervision, Anti-Money Laundering (AML) and Counter-Terrorist Financing (CFT) frameworks, and region-specific directives like the Payment Services Directive (PSD2)—have intensified their expectations for accurate, real-time, and verifiable transaction oversight. These regulations require financial institutions to maintain

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

complete visibility into the origination, transformation, and movement of data throughout its lifecycle. Consequently, data lineage and provenance have become fundamental components of compliance, fraud detection, and risk mitigation strategies.

Despite increased digitalization, ensuring accuracy, consistency, and traceability of financial data remains a challenge. Transaction information often passes through multiple processing layers, third-party systems, and legacy infrastructures, each of which may alter or replicate data in ways that are not fully transparent. This fragmentation contributes to vulnerabilities in the integrity of financial records, complicates forensic analysis, and makes it difficult to reconstruct transaction histories during audits or investigations. Ensuring tamper-evidence and trustworthiness of financial data is therefore a critical yet unmet requirement in contemporary financial monitoring systems.

1.2 Problem Statement

Although financial institutions deploy sophisticated transaction monitoring systems, these systems remain limited in their ability to provide complete, end-to-end lineage of financial data. Traditional architectures are primarily centralized, with each institution maintaining its own logs, databases, and audit trails. As a result, there is no unified mechanism to guarantee that data retains its integrity as it flows across multiple entities or systems. This lack of shared transparency creates gaps that can be exploited for financial fraud, money laundering, data manipulation, or regulatory evasion.

Furthermore, data silos remain a persistent challenge. When transaction data is stored in separate systems—often using incompatible formats, inconsistent schemas, and isolated logs—achieving comprehensive provenance becomes extremely difficult. Institutions struggle to trace the full transformation path of a transaction, leading to incomplete auditability and inconsistent reporting.

Most critically, existing monitoring infrastructures lack mechanisms to ensure tamper-evidence. Records can be modified—intentionally or inadvertently—without reliable detection, undermining the trustworthiness of audit logs and limiting their usefulness in regulatory investigations. As financial systems become more complex, the risk associated with these vulnerabilities increases significantly.

These challenges highlight the pressing need for a new approach to data lineage and provenance—one that ensures immutability, transparency, and verifiable traceability across the entire financial transaction ecosystem.

2. LITERATURE REVIEW

2.1 Data Lineage Concepts

Data lineage refers to the comprehensive tracing of data as it moves through the stages of collection, processing, transformation, storage, and usage. In financial systems, lineage provides visibility into how transaction data is generated, how it changes across different processing layers, and how it is ultimately used in compliance and reporting workflows. Core lineage mechanisms include metadata capture, audit logs, transformation histories, time-stamping, and process annotations—all of which contribute to reconstructing the data lifecycle with accuracy and reliability.

However, lineage management becomes significantly more challenging in multi-institution environments where transaction data flows across banks, regulators, payment gateways, and third-party service providers. Differences in schema, metadata standards, and logging practices often break the continuity of lineage across systems. Moreover, traditional lineage systems rely on institution-controlled logs, which may lack transparency,

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

interoperability, and tamper-evidence—limitations that undermine their usefulness in regulatory investigations and cross-border compliance checks.

2.2 Data Provenance Models

Data provenance extends the concept of lineage by providing a complete record of data origin, ownership, transformation rules, and contextual metadata. Provenance models fall broadly into two categories: process-centric provenance and data-centric provenance.

- Process-centric models document the operational steps and workflows that affect data.
- Data-centric models record the properties and transformations applied to the data itself.

Both approaches provide valuable insight, but their effectiveness depends on consistent, system-wide adoption—something rarely achieved in fragmented financial environments. Many existing systems capture provenance only within local infrastructures, resulting in incomplete or incompatible records. As financial ecosystems expand through open banking, fintech partnerships, and digital payment networks, there is a pressing need for a unified mechanism capable of maintaining verifiable provenance across multiple organizations and platforms.

2.3 Blockchain Foundations in Financial Auditing

Blockchain and Distributed Ledger Technology (DLT) have gained prominence due to their inherent characteristics of decentralization, immutability, and consensus-driven verification. In financial auditing, DLT has been explored as a mechanism for enhancing data integrity, ensuring tamper-evidence, and enabling transparent access to historical records.

Blockchain's append-only structure and cryptographic hashing make it an ideal candidate for storing audit-relevant transaction metadata. Smart contracts further automate compliance processes by embedding rules directly into the ledger. These capabilities allow auditors, regulators, and financial institutions to access synchronized, verifiable transaction histories without relying on centralized authorities. Despite these benefits, current applications of blockchain in auditing remain limited to specific use cases such as digital identity verification, trade finance, or asset tokenization; comprehensive lineage frameworks for financial monitoring remain largely unexplored.

2.4 Research Gap

Despite advancements in blockchain-enabled financial audit techniques, the literature reveals a critical gap: there is no fully developed, blockchain-governed framework capable of delivering end-to-end data lineage and provenance for financial transaction monitoring. Existing solutions either focus on partial lineage, limited provenance capture, or isolated blockchain applications that do not integrate seamlessly into multi-institution financial ecosystems.

The lack of a unified governance model leaves significant vulnerabilities in audit trails, regulatory reporting, and fraud detection. This gap highlights the need for a blockchain-governed lineage system that provides:

- Immutable transaction metadata,
- Cross-institution traceability,
- Automated compliance enforcement,
- Standardized provenance recording across financial networks.

Such a system would address longstanding challenges in financial monitoring by enabling verifiable, tamper-proof, and regulatory-aligned traceability across all stages of the transaction lifecycle.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

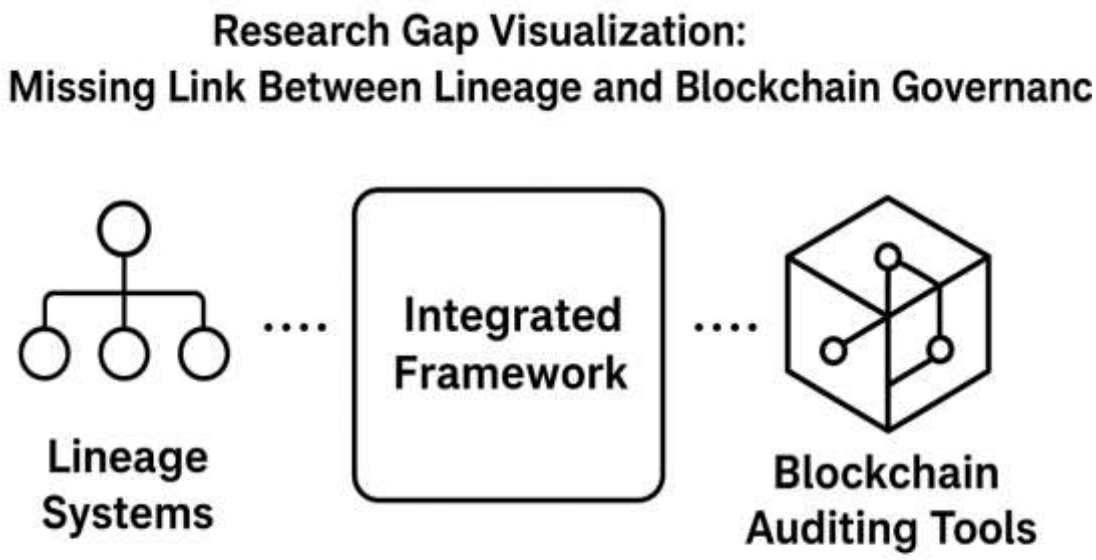

Figure: Research Gap Visualization: Missing Link Between Lineage and Blockchain Governance

3. PROPOSED BLOCKCHAIN-GOVERNED FRAMEWORK

3.1 Architecture Overview

The proposed blockchain-governed data lineage framework is designed as a multi-layer architecture that enables secure, transparent, and verifiable tracking of financial transaction histories. The architecture is composed of four distinct layers, each responsible for a critical component of the lineage and provenance lifecycle:

Data Ingestion Layer – This layer aggregates raw transaction data from various banking systems, financial applications, and transaction-processing platforms. It collects event logs, operational metadata, timestamps, and transaction attributes through standardized ingestion channels.

Lineage Encoding Layer – The collected metadata is processed, structured, and encoded into cryptographic representations. This layer constructs lineage records by combining event metadata with hashing algorithms and digital signatures.

Blockchain Ledger Layer – A permissioned blockchain serves as the immutable storage engine for encoded lineage data. Each participating institution operates a node within the consortium network, ensuring decentralized validation, synchronized consensus, and tamper-proof recordkeeping.

Regulatory Access Layer – Regulators, auditors, and compliance authorities interact with the blockchain through secure access protocols. This layer supports controlled querying, continuous audit monitoring, and automated compliance verification powered by smart contracts.

Collectively, these four layers form an integrated governance model capable of preserving end-to-end transaction lineage across multi-institution financial environments.

3.2 Lineage Capture Process

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

The lineage capture process is central to ensuring verifiable transaction histories. It consists of two primary mechanisms:

a. Transaction Metadata Extraction

Each financial event—such as a transfer, deposit, withdrawal, or settlement—generates associated metadata including timestamps, originating entity, transaction ID, processing steps, risk flags, and validation status. The system automatically extracts this metadata at every transformation point, capturing the lifecycle of the transaction in real time. This ensures consistency, completeness, and accuracy of lineage information.

b. Event Logging and Cryptographic Hashing

Once extracted, metadata is organized into structured event logs. Each event log is then transformed into a cryptographic hash, which uniquely represents the transaction state at that moment. These hashed entries are linked sequentially to form an immutable chain of provenance records. This mechanism guarantees:

- Integrity: Any alteration to the underlying data invalidates its hash.
- Non-repudiation: Each event is cryptographically tied to its previous state.
- Traceability: The full lifecycle of the transaction remains reconstructable.

The combination of automated metadata extraction and robust hashing enforces a tamper-proof lineage pipeline that is verifiable across institutional boundaries.

3.3 Smart Contract Governance

Smart contracts play a critical role in governing the flow, verification, and compliance enforcement of transaction lineage within the blockchain network.

a. Policy Enforcement

Smart contracts encode regulatory policies, compliance rules, and institutional governance standards directly into the blockchain. These rules automatically evaluate transaction lineage entries for completeness, risk indicators, and compliance violations. Any deviation from expected patterns triggers alerts or audit flags, ensuring proactive oversight.

b. Automated Compliance Reporting

Instead of relying on manual auditing or periodic reviews, smart contracts generate compliance reports automatically. The pre-defined rules detect anomalies, suspicious behaviors, or incomplete lineage records and produce real-time compliance insights for regulators. This drastically reduces audit workload, improves transparency, and enhances regulatory trust.

Together, these capabilities create an autonomous, rule-driven governance environment that guarantees consistent monitoring and reliable compliance across all participating institutions.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

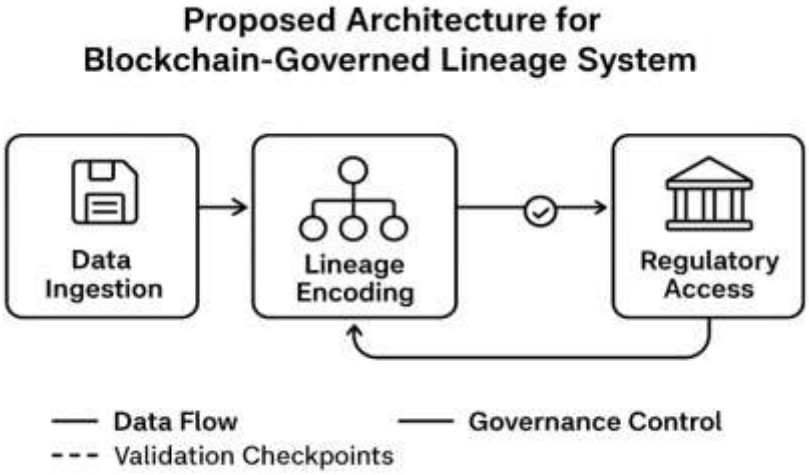

Figure 2: Proposed Architecture for Blockchain-Governed Lineage System

Table 1: Components of the Proposed Framework

Component	Description	Function
Ingestion Layer	Pulls transaction events	Metadata extraction
Provenance Engine	Generates hash + timestamp	Lineage creation
Permissioned Blockchain	Consortium of authorized nodes	Immutable ledger storage
Smart Contracts	Embedded compliance rules	Policy automation

4. SYSTEM WORKFLOW AND IMPLEMENTATION MODEL

4.1 Full Transaction Life-Cycle Flow

The proposed blockchain-governed system follows a structured workflow that ensures end-to-end visibility and verifiable provenance throughout the transaction lifecycle. The complete process can be summarized as follows:

Transaction → Metadata Extraction → Provenance Encoding → Blockchain Recording → Continuous Monitoring

Transaction Initiation:

A financial event is generated, such as a funds transfer, payment authorization, or settlement instruction. The originating system records preliminary attributes including timestamps, transaction identifiers, source/destination accounts, and risk-related metadata.

Metadata Extraction:

The ingestion layer automatically captures operational and contextual metadata from transaction logs, processing engines, and intermediary systems. This includes transformation steps, validation checks, approval states, and audit trail references.

Provenance Encoding:

Metadata is processed into structured lineage records. Each record is cryptographically hashed and linked to previous states, forming a sequence of immutable provenance entries. Digital signatures ensure authenticity and non-repudiation.

Blockchain Recording:

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

The hashed provenance entries are transmitted to the permissioned blockchain network, where they undergo consensus validation. Once validated, the entries are stored as part of a distributed, tamper-proof ledger shared across consortium participants.

Continuous Monitoring and Access:

Regulators, auditors, and financial institutions interact with the lineage records through secure API interfaces and permissioned access controls. Real-time alerts, compliance checks, and automated analytics ensure ongoing oversight.

This lifecycle enables granular traceability from the moment a transaction is created to its final audit and regulatory review.

4.2 Validation and Consensus Process

In the proposed architecture, consensus mechanisms are specifically chosen to support the requirements of financial institutions, such as high throughput, low latency, and robust fault tolerance. Since the blockchain is operated by a consortium of trusted institutions, permissioned consensus algorithms are preferred.

a. Byzantine Fault Tolerant (BFT) and PBFT Consensus

BFT-based algorithms, such as Practical Byzantine Fault Tolerance (PBFT), ensure that the network remains operational even if a portion of nodes behave maliciously or fail unexpectedly. PBFT achieves deterministic finality, making it ideal for regulatory use-cases where the immutability of lineage records must be guaranteed.

b. RAFT Consensus

RAFT provides a simpler leader-based mechanism that emphasizes performance and operational efficiency. While it does not guard against malicious nodes, it is suitable for highly trusted environments where all members are verified entities, such as banking consortiums. RAFT offers:

- Faster block finalization,
- Simplified network coordination,
- Lower computational overhead.

Both PBFT and RAFT can be adapted depending on the governance model, risk tolerance, and operational requirements of the financial ecosystem.

4.3 Security and Privacy Controls

Given the highly sensitive nature of financial transaction data, the framework integrates a multi-layered security architecture to protect confidentiality, integrity, and controlled access.

a. Access Control Lists (ACLs)

Permissioned blockchain environments enforce strict identity-based access control, ensuring that only authorized nodes or users can submit, modify, or query provenance records. ACLs restrict data visibility according to institutional roles, regulatory requirements, and data governance policies.

b. Zero-Knowledge Proofs (ZKPs)

ZKPs enable validators to confirm the correctness of a transaction or compliance status without revealing underlying sensitive information. This capability is valuable for regulatory monitoring, cross-border transactions, and inter-bank audits where privacy must be preserved while ensuring verification.

c. Encryption and Pseudonymization

All metadata transmitted to the blockchain is encrypted using industry-standard cryptographic protocols. Sensitive identifiers may be pseudonymized or tokenized to minimize exposure. This ensures that even if ledger access were compromised, the data remains unintelligible without proper authorization.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

These combined security mechanisms ensure compliance with privacy regulations such as GDPR, PCI DSS, and financial-sector cybersecurity policies.

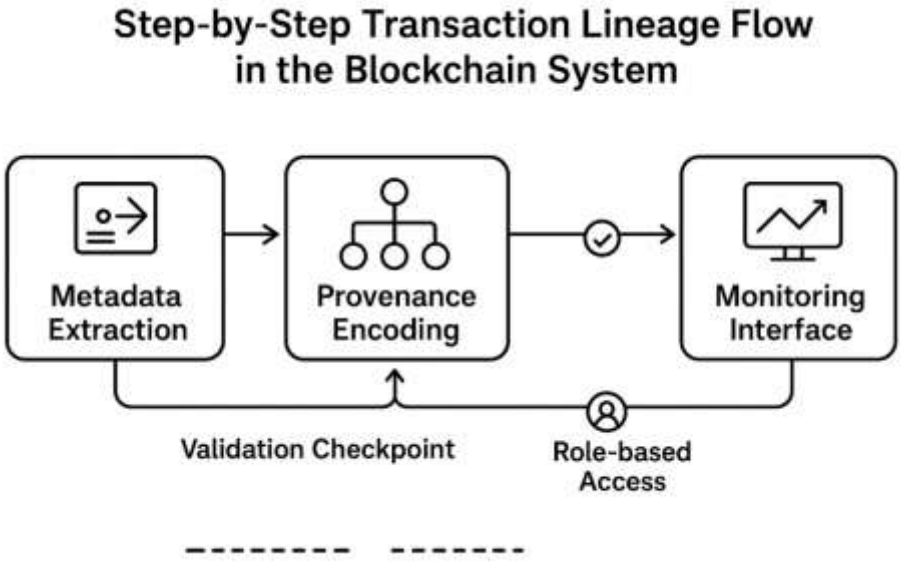

Figure 3: Step-by-Step Transaction Lineage Flow in the Blockchain System

5. EVALUATION AND DISCUSSION

5.1 Expected Outcomes

The proposed blockchain-governed data lineage framework is designed to generate several key outcomes that directly enhance financial monitoring, compliance, and auditability:

Stronger Audit Trail:

The combination of cryptographically hashed provenance records and immutable blockchain storage ensures that every transaction is traceable from initiation to final settlement. This end-to-end visibility minimizes the risk of data manipulation and provides a tamper-evident audit trail.

Real-Time Compliance:

Smart contract governance enables continuous validation of regulatory policies and compliance rules. Financial institutions and regulators can access real-time compliance insights without relying on periodic manual audits, resulting in faster detection of irregularities.

Reduced Fraud and Manipulation:

By ensuring data integrity across multiple nodes in a permissioned blockchain network, the system reduces opportunities for fraudulent activity. Unauthorized modifications are immediately detectable, and anomalies can trigger automated alerts for further investigation.

5.2 Comparative Evaluation

The blockchain-based framework offers significant improvements over traditional monitoring systems:

Enhanced Visibility:

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Unlike conventional centralized systems, which store data in siloed databases with limited cross-institution transparency, the blockchain framework enables a unified, distributed ledger accessible by all authorized participants.

Automated Verification:

Smart contracts remove reliance on manual compliance checks, reducing operational overhead while improving the accuracy and timeliness of regulatory reporting.

Tamper-Resistance:

Immutable ledger entries and cryptographic hashes ensure that provenance records cannot be altered retroactively, a key limitation in legacy systems.

A comparative analysis indicates that blockchain-based lineage significantly outperforms traditional approaches in traceability, auditability, and compliance efficiency.

5.3 Limitations

Despite its advantages, the proposed framework has several limitations that must be considered:

Scalability:

As the transaction volume grows, blockchain storage and consensus mechanisms may introduce latency and increased computational overhead. Future work should explore sharding or layer-2 solutions to mitigate scalability challenges.

Interoperability:

Integrating blockchain-based lineage systems with existing heterogeneous financial infrastructures requires standardization and mapping between legacy systems and modern DLT frameworks.

Organizational Readiness:

Adoption of blockchain governance necessitates institutional alignment, staff training, and updates to internal policies. Resistance to technological change may slow deployment and limit system effectiveness.

5.4 Future Work

Future research and development can enhance the framework's capabilities and broader adoption:

AI-Driven Anomaly Detection:

Integrating machine learning models can detect suspicious transactions, abnormal patterns, and potential fraud in real time using blockchain-stored provenance data.

Global-Coded AML Smart Contract Standards:

Developing standardized smart contract templates aligned with global AML/CFT regulations would facilitate cross-border compliance, interoperability between institutions, and streamlined regulatory reporting.

These future directions aim to further strengthen the robustness, scalability, and regulatory effectiveness of blockchain-governed financial transaction monitoring systems.

6. CONCLUSION

This study presents a blockchain-governed framework for data lineage and provenance in financial transaction monitoring, addressing critical challenges in auditability, compliance, and fraud detection. Traditional monitoring systems are limited by fragmented data storage, lack of end-to-end traceability, and susceptibility to tampering. By integrating metadata extraction, cryptographic hashing, permissioned blockchain ledgers, and smart contract-based governance, the proposed framework ensures immutable, transparent, and verifiable transaction histories across multi-institution financial ecosystems.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

The evaluation highlights several key advantages: a stronger audit trail, real-time compliance reporting, enhanced visibility compared to traditional systems, and reduced risk of fraud and data manipulation. While scalability, interoperability, and organizational readiness remain challenges, the framework establishes a foundation for more robust, transparent, and efficient financial monitoring.

Future work can further enhance the system by integrating AI-driven anomaly detection and developing globally standardized AML smart contract templates to support cross-border regulatory compliance. Overall, this research demonstrates that blockchain-governed data lineage and provenance can substantially strengthen financial risk management, regulatory adherence, and operational transparency, paving the way for next-generation financial monitoring systems.

REFERENCES

1) Auer, R., & Claessens, S. (2020). *Regulating fintech: Observations from market developments.* BIS Quarterly Review, Bank for International Settlements.
2) Bai, Y., & Lin, X. (2021). *Blockchain applications in financial services: Automated settlements and fraud prevention.* Journal of Network and Computer Applications, 177, 102952.
3) Buterin, V. (2014). *A next-generation smart contract and decentralized application platform.* Ethereum White Paper. https://ethereum.org
4) Chen, Y., Li, X., & Zhao, R. (2021). *Fraud detection and prevention in card payments using blockchain technology.* IEEE Transactions on Neural Networks and Learning Systems, 32(9), 4305–4317.
5) Gai, K., Qiu, M., & Sun, X. (2020). *Blockchain-enabled financial services: Challenges and opportunities.* Journal of Network and Computer Applications, 115, 1–11.
6) Hardjono, T., & Smith, N. (2019). *Decentralized trusted computing for financial ecosystems.* MIT Connection Science Working Paper.
7) Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). *Bitcoin and cryptocurrency technologies: A comprehensive introduction.* Princeton University Press.
8) Peters, G. W., & Panayi, E. (2016). *Understanding modern banking ledgers through blockchain technologies.* In Banking Beyond Banks and Money (pp. 239–278). Springer.
9) Rao, B., & Upadhyaya, P. (2021). *Smart contract-based automation for payment disputes and chargeback management.* International Journal of Financial Innovation, 3(2), 45–60.
10) Schär, F. (2021). *Decentralized finance: On blockchain- and smart contract-based financial markets.* Federal Reserve Bank of St. Louis Review, 103(2), 153–174.
11) Wang, S., Ouyang, L., Yuan, Y., Ni, X., Han, X., & Wang, F. (2019). *Blockchain-enabled smart contracts: Architecture, applications, and future trends.* IEEE Transactions on Systems, Man, and Cybernetics, 49(11), 2266–2277.
12) Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2018). *An overview of blockchain technology: Architecture, consensus, and future trends.* Proceedings of the 2018 IEEE International Congress on Big Data, 557–564.