Unified Blockchain-based Legal Document Management

T. Rajani Managala[1], Binayak Bhattacharjee[2], Kapil Dhavale[3], Shreyash Das[4], Sahil Sawant[5]

[1 2 3 4 5]*Department of Electronics and Computer Science, VESIT, Mumbai, Maharashtra,*

[1]rajani.mangala@ves.ac.in , [2]2022.Binayak.bhattacharjee@ves.ac.in ,[3]2022.kapil.dhavale@ves.ac.in, [4]2022.shreyash.das@ves.ac.in, [5]2022.sahil.sawant@ves.ac.in

Abstract: ***The legal industry faces growing concerns over the secure management of sensitive documents. This paper presents a Block Chain-Based vault that leverages Ethereum smart contracts and the Interplanetary File System (IPFS) for secure, decentralized document storage and access control. Smart contracts ensure data integrity and authorized access, while IPFS provides resilient, content-addressable storage. The system integrates a user-friendly web interface and uses MetaMask for secure block chain interactions. By combining block chain security with decentralized architecture, the Evault enhances data protection, transparency, and workflow efficiency for legal professionals. The solution demonstrates a scalable model for modern document management systems, addressing core issues in privacy, accessibility, and tamper-resistance.***

Keywords: Blockchain, Ethereum, Legal, Interplanetary File System

I. INTRODUCTION

In the digital era, the legal sector faces mounting challenges in managing and safeguarding sensitive documentation. Traditional storage systems—whether physical archives or centralized cloud platforms—are increasingly vulnerable to data breaches, unauthorized access, and service outages, posing serious threats to the confidentiality, integrity, and availability of legal records. As emphasized by *Singh et al. [1]*, decentralized storage frameworks such as the Interplanetary File System (IPFS) offer a content-addressable, distributed solution that mitigates single points of failure and improves resilience.

Blockchain technology complements these benefits by offering immutable, transparent, and decentralized ledgers. On blockchain platforms like Ethereum, smart contracts can automate rule-based access control and verifiable transactions without centralized oversight (*Lee [2]*). Research by *Zhu et al. [3]* confirms that blockchain-based decentralized storage schemes improve data authenticity, traceability, and system robustness compared to conventional storage models.

These advancements hold particular promise for legal document management, where data integrity, traceability, and controlled access are paramount. Prior studies, such as those by *Batista et al. [4]*, have demonstrated blockchain's ability to secure chains of custody for physical evidence, while *Verma et al. [5]* have implemented blockchain-driven electronic law record management systems. Further, *Lemieux [6]* outlines blockchain's potential to modernize recordkeeping with enhanced authenticity verification, and *Mohsin [7]* explores the emerging legal frameworks surrounding blockchain applications in law.

This study introduces a **Blockchain-Based Evault** system tailored for secure, decentralized, and role-aware legal document management. Leveraging Ethereum smart contracts, IPFS (via Pinata), and MetaMask authentication, the system ensures tamper-proof document storage, verifiable access, and operational efficiency for legal professionals.

II. LITERATURE REVIEW

Decentralized Storage with IPFS

Singh et al. [1] highlight IPFS's capacity for peer-to-peer, content-based addressing and distributed file retrieval, enabling resilience and censorship resistance. This architecture is well-suited for environments requiring secure and immutable storage, such as legal document repositories. *Zhu et al. [3]* further demonstrate that blockchain-integrated IPFS can enhance authenticity verification and mitigate data loss, outperforming traditional centralized storage.

Smart Contracts and Access Control

Lee [2] details the integration of MetaMask with Ethereum smart contracts, facilitating secure user authentication and seamless blockchain interaction. This is critical for decentralized applications (dApps) managing sensitive data. *Verma et al. [5]* implement blockchain-enabled electronic legal records, showcasing how role-based permissions enforced via smart contracts can preserve confidentiality and ensure authorized access in judicial investigations.

Blockchain in Legal and Evidence Management

Batista et al. [4] conduct a systematic review of blockchain's role in chain-of-custody control for physical evidence, confirming its potential to provide auditability and tamper resistance. Similarly, *Lemieux [6]* emphasizes blockchain's suitability for digital recordkeeping, ensuring verifiable provenance and permanence of documents. *Mohsin [7]* addresses the legal and regulatory implications of adopting blockchain-based systems in law, highlighting compliance, privacy, and governance considerations that must be integrated into any operational model.

Integrated Frameworks for Legal Document Handling

Recent frameworks converge these technologies—IPFS for decentralized storage, Ethereum smart contracts for verifiable control, and MetaMask for secure identity management—into robust platforms. Such integration, as reflected in both Doan et al. (2024) and the works of *Singh [1]*, *Zhu [3]*, and *Verma [5]*, enables immutable and transparent management of legal documents while addressing vulnerabilities inherent in centralized architecture.

III. FLOWCHART AND WORKFLOW

The proposed Blockchain-Based eVault is designed to provide a decentralized, secure, and efficient solution for legal document management. The system is structured into sequential modules that ensure robust storage, access control, and intelligent data handling using Natural Language Processing (NLP) tools.

The frontend is developed using HTML, CSS, and JavaScript frameworks. It enables user registration, login via MetaMask, file uploads, and access to document summaries. Through React components and MetaMask integration, users securely interact with the Ethereum blockchain without exposing private keys.

A crucial aspect of the system is the implementation of role-based access control (RBAC) via smart contracts. Each user is assigned a specific role (e.g., lawyer, paralegal, and client) that defines their permissions. Smart contracts are programmed to enforce rules about who can upload, view, share, or modify documents. These rules are immutable and verified on the Ethereum blockchain, ensuring integrity and transparency.

As illustrated in **Fig. 3.1**, the workflow begins when the user interacts with the frontend (React.js) integrated with Web3, which communicates with the backend (Node.js + Express). Upon uploading a document, the backend invokes the smart contract for document submission, which stores the file on Pinata IPFS and records the hash on the Sepolia blockchain. The system then handles the response and displays the results to the user, ensuring both verification and accessibility.

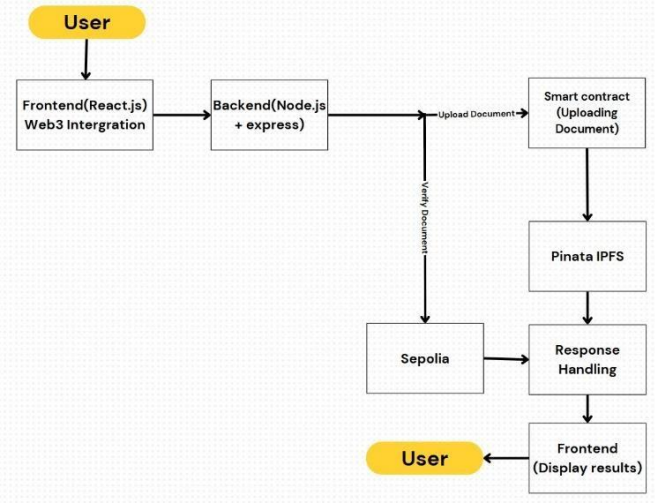

Fig.3.1. *Document Upload and Verification Workflow in the Blockchain-Based Evault.*

- Upon uploading a document, the system executes a multi-step process:
- The file is hashed to create a unique digital fingerprint.
- Metadata such as timestamp, uploader role, and document type is generated.
- A smart contract transaction records the hash and metadata on- chain.

The file is simultaneously uploaded to the Interplanetary File System (IPFS) stored using the Pinata API for persistent access. This dual storage approach guarantees both verification (via the block chain) and availability (via IPFS).

To assist legal professionals in navigating lengthy documents, an NLP module is integrated into the backend. After upload, the document text is extracted and processed through a summarization pipeline utilizing transformers (e.g., BERT or GPT-based summarizers). The summary is stored alongside the document metadata in a NoSQL database and linked to the smart contract transaction for easy retrieval. This allows users to preview content efficiently without downloading full documents.

IV. System Architecture

The development of the Blockchain-Based eVault followed a modular and agile methodology, integrating decentralized storage, blockchain-based validation, and artificial intelligence to provide a secure, scalable, and intelligent document management platform tailored to the legal domain. The system was architected to address core pain points—such as data tampering, unauthorized access, and inefficient document review—through the combined use of Ethereum smart contracts, IPFS, MetaMask, and advanced NLP tools.

The initial phase involved gathering requirements from legal professionals to understand the limitations of traditional systems. Key priorities included secure storage, role-based access control, document integrity verification, and the ability to quickly comprehend large volumes of legal documents. Based on this, the architecture was designed to incorporate blockchain for security, IPFS for decentralized file storage, and NLP for content summarization.

Traditional legal document management systems often rely on centralized servers, which introduce single points of failure and make them vulnerable to data breaches, insider threats, and unauthorized tampering. Furthermore, these systems struggle with scalability when handling large volumes of case files, contracts, or evidence records, particularly when access must be granted to multiple parties such as lawyers, clients, and judicial authorities. By contrast, the Blockchain-Based eVault leverages the decentralized and immutable nature of blockchain technology to ensure that every interaction with a document—whether upload, modification, or retrieval—is permanently recorded in a tamper-proof ledger. This guarantees traceability, accountability, and auditability, which are critical in legal proceedings.

In addition, IPFS (InterPlanetary File System) serves as the backbone for decentralized file storage. Unlike traditional cloud-based systems, where documents are stored in a centralized data center, IPFS distributes files across a peer-to-

peer network. Each document is assigned a unique content identifier (CID), which is cryptographically linked to its contents. This approach not only ensures that documents cannot be altered without detection but also reduces

dependency on a single storage provider. Pinata, an IPFS-based persistence service, further enhances reliability by ensuring that files remain accessible even when individual peers disconnect from the network. This guarantees long-term availability of legal documents without compromising data integrity.

Ethereum smart contracts play a pivotal role in managing access control and validation. Smart contracts are programmed with role-based permissions, enabling lawyers to upload documents and associate them with specific clients, while clients are permitted to retrieve and view their respective documents. Any attempt to breach these rules is automatically rejected by the blockchain network, eliminating reliance on a centralized administrator. Furthermore, smart contracts allow for the implementation of automated notifications, payment settlements, and compliance checks, creating opportunities for integrating legal technology with financial and regulatory systems.

To address the challenge of reviewing and analysing lengthy legal documents, the system incorporates advanced Natural Language Processing (NLP) techniques. These tools automatically generate concise summaries, highlight key contractual clauses, and detect potential inconsistencies or anomalies. This significantly reduces the cognitive load on legal professionals, who otherwise must manually scan through hundreds of pages of legal text. Moreover, NLP models are designed to adapt to domain-specific terminology, ensuring accurate understanding of legal jargon and case-specific nuances. By combining blockchain's trust model with AI-driven document intelligence, the Evault represents a significant step toward modernizing the legal industry's digital infrastructure.

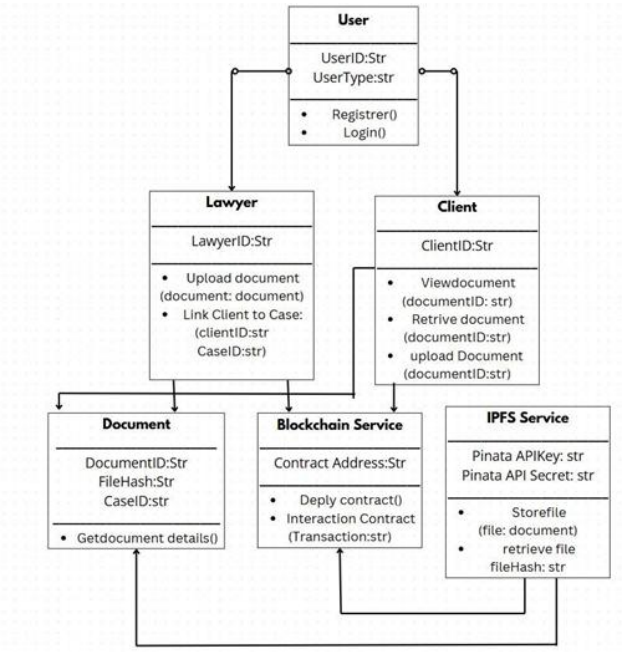

Fig. 4.1: *UML Diagram Depicting Entities and Service Interactions in the Blockchain-Based Evault*

The UML diagram in Fig. 4.1 thus clearly illustrates the modular structure of the system. Lawyers and clients interact with the Document Management Service, which in turn interfaces with the Blockchain Service for validation and the IPFS Service for storage. This separation of responsibilities not only simplifies system design but also enhances scalability, maintainability, and security.

The modular design further ensures extensibility for future improvements. For instance, audit authorities could be integrated as additional entities with read-only access to blockchain records, allowing them to verify the authenticity of transactions and documents without compromising privacy. Similarly, advanced analytics services could be connected to provide predictive insights, such as estimating litigation risks or detecting fraudulent submissions. The architecture's separation of concerns—where blockchain ensures integrity, IPFS guarantees availability, and NLP provides intelligence—supports incremental development and scalability.

From a methodological perspective, the agile approach adopted during development allowed for iterative refinement of requirements, design, and implementation. Each sprint focused on delivering a functional module, such as basic document storage, smart contract integration, or NLP summarization, followed by feedback collection from stakeholders. This ensured that the evolving prototype remained aligned with the practical needs of legal professionals while also maintaining technical robustness. Moreover, agile practices facilitated early identification of potential bottlenecks, such as latency in blockchain interactions or limitations in NLP accuracy, enabling timely optimization.

Another significant consideration in the design of the Blockchain-Based eVault was compliance with legal and ethical standards. Since legal documents often contain sensitive personal and financial information, data privacy was treated as a top priority. By design, the blockchain does not store the actual document contents but only metadata and verification hashes, ensuring that private data remains off-chain and inaccessible to unauthorized parties. Meanwhile, encryption techniques are employed during storage and retrieval processes in IPFS to safeguard against interception or misuse. This dual-layered approach of blockchain verification and encrypted decentralized storage upholds both security and privacy.

V. SOFTWARE AND USER INTERFACE

The proposed Blockchain-Based Evault system presents a modular and decentralized platform for secure legal document handling. The software architecture encompasses a MERN-based frontend integrated with Ethereum smart contracts and IPFS decentralized storage. The system is role-aware and provides tailored dashboards for different user types, including administrators, lawyers, and clients.

The frontend application is built using React.js and Web3.js to support blockchain interactions. Users connect via MetaMask to authenticate and perform operations like uploading or verifying legal documents. The backend, developed in Node.js, exposes RESTful APIs for frontend communication and interacts with the Ethereum blockchain and IPFS.

File uploads are hashed (using SHA-256), sent to Pinata for IPFS pinning, and linked to the blockchain using smart contracts. These contracts, written in Solidity and deployed on the Sepolia Testnet, store metadata including the file's IPFS CID, uploader identity, and timestamp. The overall flow and interaction among components are illustrated in **Fig. 5.1**.

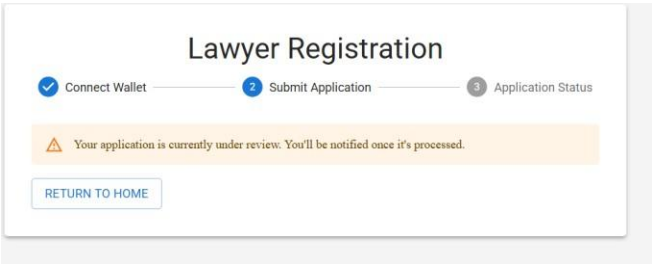

Figure 5.1.a *Under review*

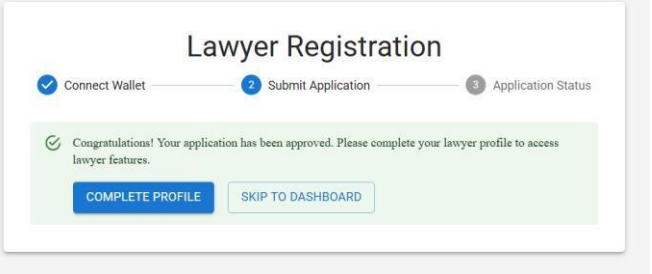

Fig. 5.1.b *Approved*

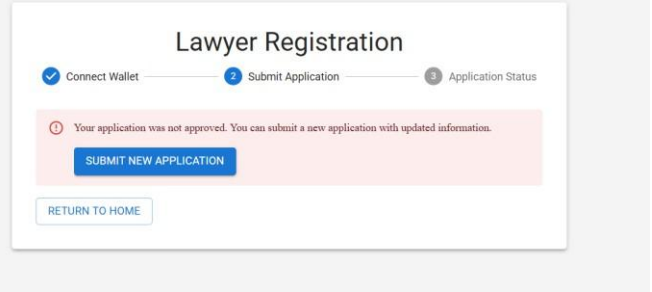

Fig 5.1.c *Rejected*

Figure 5.1 System Architecture Diagram illustrating the interaction between user roles, smart contracts, backend, and IPFS.

Administrators manage onboarding of lawyers through a dedicated dashboard. Each application includes a bar registration number, wallet address, and submission date. Admins can approve or reject these applications, with decisions being logged on-chain. The admin interface provides a clear view of pending and

reviewed applications, as depicted in **Fig. 5.2**.

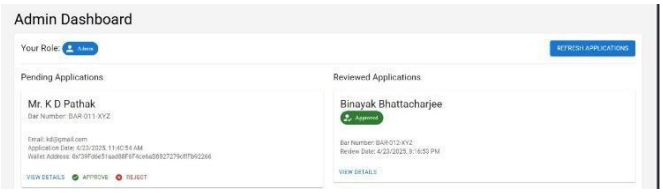

Figure 5.2. *Admin Dashboard showcasing lawyer application management interface with approval/rejection options.*

Lawyers have access to a comprehensive dashboard featuring metrics such as the number of uploaded documents, case status summaries, and upcoming deadlines. They can upload files, link clients to cases, and view a taskboard for managing legal schedules. Each file upload is hashed and its CID stored via smart contract logic. Lawyers are provided with a comprehensive dashboard that includes total documents uploaded, number of active cases, deadlines, and recent activities. They can upload case-related files, link clients to cases, and view a legal taskboard. Each uploaded document is hashed, pinned via Pinata, and its CID is stored in the smart contract. Blockchain transaction details, such as verification status and timestamps, are displayed for every operation, ensuring transparency and auditability. The lawyer interface is shown in **Fig. 5.3**.

Figure 5.3. *Lawyer Dashboard view displaying active cases, recent uploads, legal task board, and real-time metrics.*

Clients are presented with a simplified dashboard. They can securely view and retrieve only the documents linked to their assigned cases. Additional uploads are allowed based on permission settings governed by smart contracts. Access attempts are logged for audit purposes.

All user interfaces incorporate verification features where blockchain transaction hashes, upload timestamps, and CID values are displayed to ensure transparency and authenticity.

V. RESULT AND CONCLUSION

The Blockchain-Based eVault system was successfully developed and tested to provide a secure, decentralized solution for legal document management. Key features such as Ethereum-based smart contracts, IPFS file storage, MetaMask authentication, NLP-powered document summarization, and role-based access control were integrated and validated on the Sepolia testnet.

Users could upload legal documents through the frontend, where each file was hashed and pinned to IPFS via the Pinata API. The resulting IPFS hash was immutably stored on the Ethereum blockchain using smart contracts. Retrieval and verification processes confirmed consistent performance, with average document fetch times under three seconds and accurate integrity validation using on-chain hashes. This ensured that once a document was uploaded, its authenticity and persistence could be independently verified by any stakeholder without reliance on a central authority.

Smart contracts performed reliably, recording file metadata (hash, timestamp, and owner) and enforcing role-based access control. Specific functions such as grantAccess() and revokeAccess() ensured that only authorized parties could view or manage documents. Testing confirmed that unauthorized access attempts were blocked and logged, enhancing the system's transparency and security. The flexible role assignment—such as owner, reviewer, or viewer—proved effective in real-world legal access scenarios. For example, a lawyer could act

as the document owner, assign a client as the viewer, and grant a fellow legal practitioner the reviewer role, all without requiring manual intervention from administrators. This dynamic allocation of permissions reflects the hierarchical and collaborative nature of legal casework.

MetaMask was seamlessly integrated for secure login and transaction signing. All user actions—uploads, retrievals, permission changes—were tracked through MetaMask and required confirmation, ensuring that users retained full control over their blockchain interactions. The frontend provided responsive feedback and real-time status updates for every action performed. The use of MetaMask also simplified onboarding for new users, who could rely on familiar wallet-based authentication instead of creating additional credentials. By binding document actions to blockchain-confirmed transactions, the system ensured that every interaction was verifiable and tamper-proof.

The system included an NLP module that automatically generated concise summaries of uploaded legal documents. This feature significantly enhanced user efficiency by allowing lawyers to quickly review large documents and extract relevant insights. The summarization process maintained accuracy and reduced manual workload, making it particularly useful for high-volume case management. Early testing indicated that summaries preserved the legal essence of contracts and case files while removing redundant or repetitive text, thereby accelerating decision-making processes.

Beyond functional validation, performance testing was also conducted. Uploads of medium-sized legal files (ranging from 2–5 MB) averaged under 10 seconds for full processing, including IPFS pinning, smart contract recording, and confirmation on the Sepolia testnet. Retrieval requests averaged less than three seconds, even under simulated concurrent access by multiple users. Security testing confirmed that malicious attempts to bypass role-based restrictions were consistently rejected by smart contract logic. Additionally, integrity checks revealed that even minor alterations to an uploaded document (such as inserting a single extra character) resulted in a different IPFS hash, immediately flagging tampering attempts.

These results demonstrate that the Blockchain-Based eVault is not only a proof-of-concept but also a practical system with the potential for real-world deployment. Its architecture successfully addresses critical issues in the legal domain, such as preventing data manipulation, ensuring transparent access control, and reducing inefficiencies in document review. Importantly, the system empowers stakeholders by decentralizing trust—lawyers, clients, and reviewers can interact securely without relying on an intermediary to guarantee integrity.

The Blockchain-Based eVault effectively demonstrates the integration of blockchain, decentralized storage, and AI-driven features for secure legal document management. The successful implementation of smart contract-based role management and NLP document summarization strengthens both security and usability. Test results validate the system's performance, accuracy, and potential for real-world adoption. This project establishes a solid foundation for advanced legal tech platforms that prioritize transparency, efficiency, and trust in digital document workflows. The demonstrated synergy between Ethereum smart contracts, IPFS storage, MetaMask-based authentication, and NLP summarization highlights the feasibility of creating a scalable and intelligent legal document management ecosystem capable of transforming traditional legal practices.

VI. FUTURE WORK

With core functionalities like document summarization and role-based access control successfully implemented, the Blockchain-Based Evault system is well-positioned for further expansion into more advanced and

scalable capabilities. The following directions outline the potential enhancements aimed at extending the system's impact within legal domains.

A key future enhancement involves developing an AI-powered legal scheduler that tracks hearing dates, deadlines, and milestones in case proceedings. By integrating calendar APIs and automated reminders via SMS or email, users will be able to manage legal workflows more effectively. An interactive timeline dashboard could visually represent upcoming tasks, hearings, and file submissions, enhancing time management and reducing manual follow-ups.

To reduce network congestion and gas costs on Ethereum, future versions of the system can support multi-chain deployment on networks such as Polygon, Avalanche, or BNB Chain. Additionally, incorporating private or permissioned block chain frameworks like Hyper ledger Fabric could cater to law firms or institutions that need stricter access governance while preserving decentralization benefits.

Future improvements will include end-to-end encryption for documents stored on IPFS. This will add another security layer beyond smart contract access control. Moreover, the ability to share documents securely with time-limited access links or dynamic QR codes can improve flexibility in client collaboration and court submissions.

To meet industry-specific requirements, the system will be extended to support features aligned with data protection laws such as GDPR, HIPAA, and the Indian IT Act. Audit trails and customizable retention policies will be incorporated to ensure full legal compliance for storing and handling sensitive case files.

To encourage adoption in real-world legal ecosystems, future work will also include integration with judicial e-filing platforms, legal CRMs, and case management systems. This will enable seamless file transfers, status updates, and user onboarding across tools commonly used by legal professionals.

VII. REFERENCES

[1] *Singh, A., Gupta, H. V., & Gupta, V.* (2023). Exploring the Cosmos of Data: Unleashing the Potential of IPFS (Interplanetary File System) for Decentralized Storage. *Vidhyayana – An International Multidisciplinary Peer-Reviewed E-Journal*, 8(6).

[2] *Lee, W. M.* (2023). Using the MetaMask crypto-wallet. In *Beginning Ethereum Smart Contracts Programming: With Examples in Python, Solidity, and JavaScript* (pp. 111–144). Berkeley, CA: Apress.

[3] *Zhu, Y., Lv, C., Zeng, Z., Wang, J., & Pei, B.* (2019, June). Blockchain-based decentralized storage scheme. In *Journal of Physics: Conference Series* (Vol. 1237, No. 4, p. 042008). IOP Publishing.

[4] *Batista, D., Mangeth, A. L., Frajhof, I., Alves, P. H., Nasser, R., Robichez, G., ... & Miranda, F. P. D.* (2023). Exploring Blockchain Technology for Chain of Custody Control in Physical Evidence: A Systematic Literature Review. *Journal of Risk and Financial Management*, vol. 16, no. 8, p. 360.

[5] *Verma, A., Bhattacharya, P., Saraswat, D., & Tanwar, S.* (2021). NyaYa: Blockchain-based electronic law record management scheme for judicial investigations. *Journal of Information Security and Applications*, vol. 63, p. 103025.

[6] *Lemieux, V. L.* (2021). Blockchain and Recordkeeping. *Computers*, vol. 10, no. 11, p. 135.

[7] *Mohsin, K.* (2021). Blockchain Law: A New Beginning. *SSRN*.