

Retrieval-Augmented Generation (RAG) Framework for Digital Libraries Utilizing Blockchain-Based Document Hashing

Jomy George

PG Scholar

*Department of Computer Applications
Amal Jyothi College of Engineering, Autonomous
Kanjirappally, Kottayam, Kerala, India
jomygeorge2026@mca.ajce.in*

Gloriya Mathew

Assistant Professor

*Department of Computer Applications
Amal Jyothi College of Engineering, Autonomous
Kanjirappally, Kottayam, Kerala, India
gloriyamathew@amaljyothi.ac.in*

Abstract- The rapid advancement of Retrieval-Augmented Generation (RAG) has transformed digital libraries by enabling sophisticated semantic interactions with literary works, yet these systems are entirely dependent on the integrity of their source material. As digital publishing grows in scale, the risk of malicious document manipulation expands with it, making content provenance a critical and largely overlooked challenge. Any unauthorized tampering or "source-level poisoning" can lead the AI to generate misinformation and hallucinations, quietly eroding trust between authors and readers — often without leaving any visible trace. This research proposes a framework that integrates blockchain-based document hashing to establish a verifiable "root of trust" for RAG-powered systems. By utilizing SHA-256 cryptographic hashing to notarize the original state of a literary asset on an immutable ledger, the system performs real-time integrity checks before ingestion and retrieval. The architecture leverages semantic chunking and high-dimensional vector embeddings for accurate knowledge retrieval through a persistent vector database, strictly gated by on-chain verification. Implementation results demonstrate that the framework reliably identifies tampered assets and prevents the RAG engine from acting on unverified content. Ultimately, this ensures that AI-generated summaries and insights remain faithful to the original author's voice, fostering a secure and accountable digital publishing ecosystem.

Keywords: *Retrieval-Augmented Generation (RAG), Blockchain, Document Hashing, Source Integrity, Digital Libraries, Immutable Ledger, Data Provenance, Content Authenticity*

I. INTRODUCTION

In the contemporary digital era, the publishing industry is undergoing a significant transformation driven by the integration of Generative Artificial Intelligence and advanced information retrieval. Retrieval-Augmented Generation (RAG) has emerged as a technology for digital libraries, allowing Large Language Models (LLMs) to provide precise, context-aware answers by querying external datasets rather than relying solely on pre-trained internal knowledge. However, this reliance on external data introduces a critical vulnerability: the integrity of the source material. Standard digital library architectures lack a decentralized "root of trust," making them susceptible to "source-level poisoning"—a scenario where literary assets are subtly tampered with or modified. When a RAG engine processes these altered files, it inadvertently propagates misinformation and

artificial hallucinations, thereby undermining the credibility of the platform and the intellectual property rights of authors.

To address these challenges, this research presents a secure framework designed to establish absolute provenance and authenticity in digital literary marketplaces. The proposed system integrates blockchain-based document hashing with a high-performance RAG pipeline to ensure that the AI only interacts with verified, untampered content. Upon the upload of a literary work, the system generates a unique SHA-256 cryptographic hash which is then notarized on an immutable blockchain ledger, creating a tamper-evident fingerprint that is permanently tied to the document's original state. This notarized record serves as a permanent reference point for real-time integrity checks during the ingestion phase, where text is semantically chunked and stored in a persistent vector database. Any subsequent attempt to modify the source document is immediately detected when the recomputed hash fails to match the on-chain record, effectively blocking the compromised content from entering the retrieval pipeline. The retrieval layer then operates exclusively on verified chunks, ensuring that every response generated by the system can be traced back to an authenticated source. By utilizing a localized 5.65GB Ollama model for generation, the framework ensures data privacy and operational resilience, eliminating dependence on external APIs and reducing exposure to third-party vulnerabilities. The ultimate goal of this framework is to bridge the gap between AI-driven accessibility and document security, providing a transparent and trustworthy ecosystem where digital assets are both intelligent and verifiably authentic.

II. LITERATURE REVIEW

A. Evolutions in Retrieval-Augmented Generation:

The emergence of Retrieval-Augmented Generation (RAG) has addressed several fundamental limitations of standalone Large Language Models (LLMs), specifically regarding their inherent knowledge cutoffs and the tendency to generate factually incorrect information, known as hallucinations. Recent research highlights that RAG systems significantly improve the efficiency of AI models by allowing them to access external, domain-specific knowledge bases without the computational overhead of continuous fine-tuning. The integration of transformer-based sentence embeddings, such as the all-MiniLM-L6-v2 model, has further refined this process by

enabling high-dimensional semantic similarity analysis between user queries and document excerpts. However, most current literature focuses on improving retrieval speed or context relevance through hybrid architectures and persistent vector stores like ChromaDB. There remains a notable gap in research regarding the security of the underlying data; traditional RAG frameworks typically assume that the retrieved external data is authentic and untampered, leaving them vulnerable to integrity-based attacks.

B. Data Integrity and the Threat of Model Poisoning:

In the context of digital libraries and automated knowledge retrieval, "source-level poisoning" has become a critical security concern. Unlike traditional cyberattacks that target system availability, integrity attacks focus on subtly modifying the content within a knowledge base to manipulate the output of an AI model. Literature in AI safety emphasizes the need for "grounded intelligence," where an AI's responses are strictly confined to a verified context. However, as digital assets are increasingly redistributed and hosted on diverse platforms, ensuring that a file remains in its original, author-verified state is difficult. Current security paradigms often rely on centralized authorities or encrypted file formats, which can be bypassed or lack the transparency needed for public-facing digital libraries. The research community has identified a pressing need for a decentralized method to ensure that the "root of trust" for a RAG engine remains immutable throughout the document's lifecycle.

C. Blockchain-Based Notarization and Provenance:

Blockchain technology, characterized by its decentralized and immutable ledger, offers a robust solution for establishing document provenance and verification. Previous studies have explored the use of "Notary Vaults" to secure document authentication through cryptographic hashing, where a unique fingerprint of a file is stored on-chain to prevent future tampering. In a digital library ecosystem, this notarization provides a permanent, timestamped record of an author's original work. By anchoring a document's SHA-256 hash on a blockchain, any subsequent modification—no matter how minor—results in a hash mismatch, thereby alerting the system to a breach in integrity. While blockchain has been widely used for financial transactions and basic file notarization, its application as a real-time verification gate for AI retrieval represents a novel integration. This research bridges this gap by proposing a framework where the RAG engine's processing is directly contingent upon a successful blockchain-verified integrity check, ensuring that AI-driven insights are based only on authentic literary assets.

III. METHODOLOGY

The proposed framework utilizes a decoupled architecture that synchronizes decentralized document notarization with an intelligent retrieval pipeline to ensure both data integrity and grounded AI responses. The methodology is structured into two primary phases: (A) Blockchain-Based Integrity Verification and (B) The Retrieval-Augmented Generation (RAG) Pipeline.

A. Blockchain-Based Integrity Verification

The first phase focuses on establishing a "root of trust" for every literary asset uploaded to the platform. When an author submits a document in EPUB format, the system immediately generates a unique cryptographic fingerprint using the SHA-256 hashing algorithm. This hash serves as an immutable identifier for the specific version of the work. The service then utilizes the ethers.js library to notarize this hash onto a decentralized ledger. Rather than storing the entire book on the blockchain, only the 256-bit hash is recorded, ensuring a lightweight and cost-effective verification process. Before the RAG engine processes any user query, the system re-calculates the file's hash and compares it against the on-chain record. If a discrepancy is detected, the system identifies the content as tampered and halts the retrieval process to prevent the propagation of misinformation.

B. Retrieval-Augmented Generation (RAG) Pipeline

Once a document's integrity is verified, it enters the RAG pipeline, which is designed to provide high-fidelity answers without requiring the AI to be retrained on the specific content.

1. Document Ingestion and Semantic Chunking: The verified EPUB file is parsed to extract raw text, which is then partitioned into overlapping semantic chunks of 800 characters with a 200-character overlap. This overlap ensures that contextual information at the boundaries of chunks is preserved for the AI.

2. Vector Embedding and Storage: These chunks are converted into high-dimensional vector representations using the all-MiniLM-L6-v2 transformer model. These embeddings are stored in a persistent ChromaDB vector store, indexed by a unique collection name derived from the book's identifier.

3. Contextual Retrieval and Generation: When a user submits a query, it is embedded into the same vector space, and a similarity search is performed to retrieve the top 8 most relevant excerpts (Top-K=8) based on cosine distance. These verified excerpts are injected into a specialized prompt that instructs the LLM to function as a helpful book companion. The final response is generated using a localized 5.65GB Llama 3.2 model via the Ollama framework, ensuring that the process remains private, secure, and grounded strictly in the authenticated text.

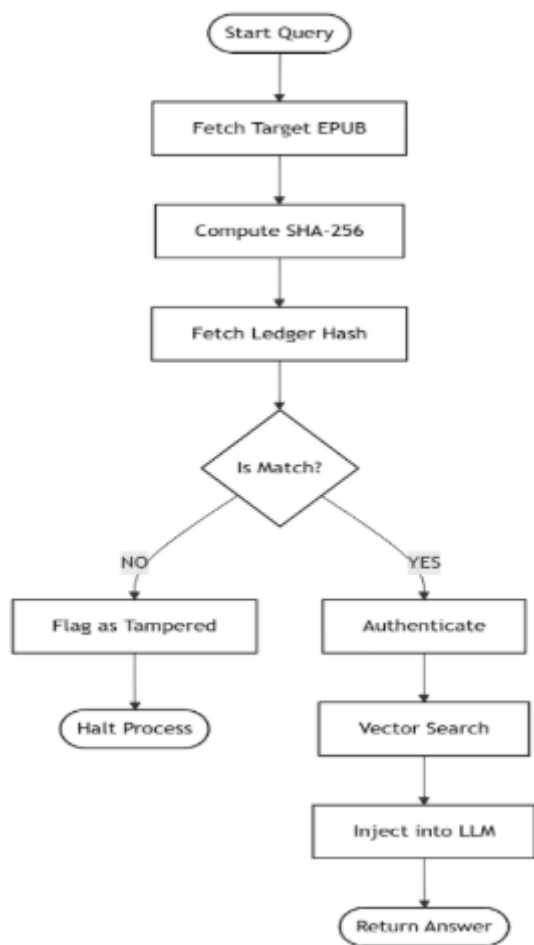


Fig. 1.A flowchart illustrating a blockchain-secured RAG framework for digital libraries. The process ensures document integrity by computing a SHA-256 hash of a target EPUB and verifying it against a ledger hash before proceeding with vector search and LLM injection.

IV. RESULTS AND DISCUSSION

The prototype implementation of the RAG framework for digital libraries successfully demonstrates the synergy between decentralized integrity verification and localized AI intelligence. The results are evaluated based on three core metrics: cryptographic notarization reliability, ingestion efficiency, and the accuracy of the grounded retrieval system.

A. Integrity Verification Performance

The system successfully implemented a "Verify-Before-Retrieve" protocol. During testing, the blockchain service module consistently generated unique SHA-256 hashes for diverse .epub files and successfully notarized them on the blockchain ledger. In scenarios where a document was subtly modified (even by a single character), the real-time hash comparison correctly identified a mismatch against the on-chain record, triggering a system-wide block on the RAG engine. This confirms that the framework effectively mitigates source-level poisoning by ensuring that only authentic literary assets are used for AI generation.

B. RAG Ingestion and Vector Search

The ingestion pipeline demonstrated high efficiency in converting raw literary content into searchable intelligence. Utilizing the all-MiniLM-L6-v2 embedding model, the system successfully partitioned documents into 800-character chunks with a 200-character overlap, ensuring context preservation. These chunks were indexed in a persistent ChromaDB store, allowing for sub-second semantic retrieval of the Top-K relevant excerpts. The use of a persistent vector client ensures that once a book is verifiably ingested, it remains available for instant querying without redundant processing.

V. FUTURE SCOPE

The current implementation of the framework establishes a solid foundation for verifiable intelligence in digital libraries, but several avenues exist for further research and technical expansion. One primary area for future development is the transition from localized or centralized storage to a Decentralized Storage Network such as IPFS (InterPlanetary File System) or Arweave. Integrating decentralized storage would ensure that the literary assets themselves are as immutable and censorship-resistant as their on-chain hashes, creating a fully decentralized content lifecycle from storage to retrieval.

Another significant enhancement involves the implementation of Smart Contract-Based Access Control. While the current system focuses on source integrity through document hashing, future iterations could utilize Non-Fungible Tokens (NFTs) to represent book ownership. This would allow the RAG engine to perform "Ownership-Gated Retrieval," where the AI only provides detailed insights or summaries if the user's cryptographic wallet contains the specific token associated with that book ID. This would create a robust, transparent royalty system where authors are directly compensated for every AI interaction grounded in their work.

From an artificial intelligence perspective, the framework could be expanded to support Multimodal Retrieval-Augmented Generation. Currently, the system extracts and processes raw text from EPUB files; however, many literary and technical works rely heavily on images, charts, and mathematical notations. Developing a multimodal RAG pipeline would enable the Ollama model to interpret and explain visual data, providing a more comprehensive companion experience. Finally, exploring Zero-Knowledge Proofs (ZK-Proofs) could allow users to verify the integrity of a document or their right to access it without revealing sensitive personal information, further enhancing the privacy-centric nature of the digital library.

VI. CONCLUSION

This research successfully presented and implemented a Retrieval-Augmented Generation (RAG) framework tailored for digital libraries, uniquely fortified with blockchain-based document hashing to ensure source integrity. By addressing the critical vulnerability of "source-level poisoning," the system establishes a decentralized root of trust that prevents the AI from propagating misinformation derived from tampered literary

assets. The integration of SHA-256 cryptographic fingerprints notarized on a decentralized ledger provides an immutable reference point, ensuring that every document processed by the RAG engine is verifiably authentic.

The technical implementation demonstrates that high-performance AI interactions can be achieved without compromising data privacy or author rights. By utilizing a localized 5.65GB Llama 3.2 model via the Ollama framework, the system provides a resilient and private environment for knowledge retrieval that does not depend on external cloud-based APIs. The use of semantic chunking and persistent vector storage through ChromaDB allows for efficient, contextually accurate responses that are strictly grounded in the verified source text.

In conclusion, the framework offers a scalable solution for the modern digital publishing industry, where trust and authenticity are paramount. By linking the intelligence of RAG with the immutability of blockchain, this research provides a blueprint for secure digital libraries that protect intellectual property while delivering advanced AI-driven user experiences. This "trust-by-design" architecture ensures that digital libraries remain a reliable source of knowledge in an era increasingly influenced by generative artificial intelligence.

and decentralized ledger concepts underpinning the integrity verification layer).

REFERENCES

- [1] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 9459-9474, 2020. (Foundational paper for the RAG architecture).
- [2] A. I. Dubey et al., "The Llama 3 Herd of Models," arXiv preprint arXiv:2407.21783, 2024. (Cites the underlying architecture for the localized 5.65GB model utilized via Ollama).
- [3] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 3982-3992, 2019. (Academic citation for the all-MiniLM-L6-v2 embedding logic).
- [4] W. Wang et al., "MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 5776-5788, 2020. (Provides technical background for fast, lightweight vector embeddings).
- [5] Ollama, "Ollama – Run Large Language Models Locally," 2024. [Online]. Available: <https://ollama.com> (Official documentation for the Ollama framework used to deploy the localized Llama 3.2 model).
- [6] G. Wood, "Ethereum: A Secure Decentralised Generalised Transaction Ledger," *Ethereum Project Yellow Paper*, vol. 151, pp. 1-32, 2014. (Foundational citation for the blockchain network and smart contract architecture used in the integrity checks).
- [7] Y. Gao et al., "Retrieval-Augmented Generation for Large Language Models: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, 2023. (Provides a comprehensive review of the modern RAG landscape).
- [8] S. Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System*, Self-published, 2008. [Online]. Available: <https://bitcoin.org/bitcoin.pdf> (Foundational reference for blockchain

DOI: 10.5281/zenodo.19542080

ISBN: 978-93-342-7372-4@2026 Amal Jyothi College of Engineering, Kanjirappally, Kottayam