



Blockchain Technology for Digital Preservation in Libraries: Case Studies and Scalability Challenges

M Hema Sundara Rao¹ and G. Siva Prasad²

¹M. Hema Sundara Rao
Research Scholar (UGC-JRF),
Department of Library and Information Science,
Andhra University, Visakhapatnam, India
E-mail: hemasundarmandala@gmail.com

²G. Siva Prasad
Professor
Department of Library and Information Science,
Andhra University, Visakhapatnam-530 003, India
E-mail: gsprasad1967@gmail.com

Abstract

Blockchain technology offers promising solutions for the digital preservation of library records, ensuring digital objects' authenticity, integrity, and provenance. This paper explores the use of blockchain in various pilot projects and case studies from institutions like the UK National Archives, MIT Media Lab, and the Internet Archive. It highlights the effectiveness and challenges in scalability, energy consumption, and cost these projects faced. Through these case studies, the paper assesses blockchain's potential for large-scale digital preservation and outlines future research directions to overcome existing barriers.

Keywords: Blockchain Technology, Digital Archives, Libraries, Case Studies, Pilot Projects, Media Lab

1. Introduction

Blockchain technology is considered as a segment of the fourth industrial revolution. Blockchain is the major technological aspect behind cryptocurrency based present day products. Blockchain technology has the capability of transforming the current Internet from “The Internet of Information Sharing” to “The Internet of Value Exchange.” Library professionals and users are not untouched with the technological changes, in fact, they are highly influenced by these changes.

Digital preservation is a critical responsibility for libraries, archives, and cultural institutions. With the rapid expansion of born-digital content, traditional methods of preserving records face significant challenges in ensuring the integrity and authenticity of information over time. Blockchain technology, known for its decentralized and immutable nature, has emerged as a potential tool to address these challenges. This paper investigates

how blockchain has been applied to digital preservation efforts through various pilot projects and case studies, assessing both its effectiveness and scalability for future use in libraries.

2. Objectives

- To examine the role of Blockchain technology in digital preservation.
- To present case Studies on the usage of Blockchain technology in digital preservation.
- To discuss challenges in Blockchain scalability.
- To assess Blockchain's potential for libraries.
- To propose future research directions.

3. Blockchain and Digital Preservation: A Brief Overview

Blockchain, initially created to power cryptocurrencies, has found applications beyond finance due to its core attributes of decentralization, transparency, immutability, and cryptographic security. These characteristics make it particularly appealing in the realm of digital preservation, where it can address critical challenges such as ensuring the authenticity, integrity, and provenance of digital assets.

The essential features of blockchain for digital preservation include:

- **Immutability:** Once data is entered, it cannot be altered without detection, providing strong safeguards against tampering.
- **Decentralization:** By eliminating a single point of failure, blockchain increases the resilience of stored records.
- **Transparency:** Blockchain's open and verifiable ledger allows full visibility into the history and records changes, promoting traceability.



Figure 1: Representing the integration of blockchain and digital preservation in libraries

Libraries and archival institutions are increasingly exploring blockchain to guard against unauthorized alterations, ensure a transparent history of modifications, and decentralize storage systems, which helps mitigate the risk of institutional data loss or failures.

4. Case Studies and Pilot Projects

Several institutions have initiated pilot projects to test blockchain's viability for digital preservation. This section presents five key case studies.

4.1. Archangel Project (UK National Archives)

The ARCHANGEL project, led by the University of Surrey in collaboration with the UK National Archives, aimed to preserve the integrity of digital records. The project used a permissioned blockchain to store cryptographic hashes of digital documents, ensuring that any document changes would be easily detectable. ARCHANGEL effectively ensured the immutability and authenticity of archives while reducing the risks of data tampering.

ARCHANGEL is a project using AI and Blockchain to secure the integrity of digital public archives, addressing the vulnerability of digital data to tampering and corruption. By employing Blockchain, a decentralized and tamper-proof ledger, ARCHANGEL stores digital fingerprints (unique identifiers) of documents. These fingerprints can be verified later to ensure document integrity, regardless of format changes over time.

However, the project faced scalability challenges, particularly regarding how much metadata could be stored on-chain. The permissioned nature of the blockchain helped mitigate some concerns about performance, as only authorized institutions participated in the network.

4.2. MIT Media Lab – Blockcerts

The MIT Media Lab, in collaboration with Learning Machine, developed Blockcerts, an open-source platform for storing academic credentials on the blockchain. This project demonstrated the potential of blockchain for digital preservation, particularly in providing secure, verifiable, and tamper-proof academic records.

Although Blockcerts proved to be highly effective for preserving small, high-value digital objects such as academic credentials, it showed that larger collections of digital objects would require off-chain storage to remain scalable.

4.3. San Jose State University – Blockchain in Libraries Initiative

San Jose State University piloted a blockchain project focused on preserving digital learning materials and student records. The project also explored blockchain's potential for decentralized cataloging and resource-sharing among libraries.

While the project confirmed that blockchain could ensure the authenticity of educational resources, it encountered challenges related to blockchain operations' cost and energy consumption, especially for libraries with constrained budgets.

4.4. Preserv Project (University of Glasgow and UK National Archives)

The Preserv Project focused on using blockchain to create immutable records for government documents. Smart contracts were used to automate the preservation process, ensuring that records were not altered without proper authorization.

This project highlighted blockchain's potential to provide transparency and tamper-proof audit trails. However, it also raised legal and ethical concerns regarding privacy and the potential conflict with the "right to be forgotten" under data protection regulations.

4.5. Project Alexandria (Wayback Machine and Blockchain)

Project Alexandria by the Internet Archive's Wayback Machine explored blockchain for web content preservation. The project stored cryptographic proofs (hashes) of archived web pages on the blockchain, allowing users to verify the authenticity of web content.

While Project Alexandria showed that blockchain could be used to enhance the trustworthiness of web archives, it struggled with the large volume of data inherent in web archiving. The blockchain was used primarily to store small amounts of metadata, while the web pages were stored off-chain.

5. Key Findings and Challenges

The case studies above demonstrate the following benefits of using blockchain for digital preservation:

- **Immutability:** Blockchain ensures that records cannot be altered without detection, offering strong protection against tampering.
- **Transparency:** Blockchain provides a complete and transparent history of modifications and access, which is useful for audit trails.
- **Decentralization:** Distributed blockchain networks reduce the risks associated with centralized storage systems.

However, significant challenges remain:

- **Scalability:** Blockchain's current structure is not optimized for storing large digital objects, making off-chain solutions necessary.
- **Cost and Energy Consumption:** Blockchain, especially Proof of Work (PoW) systems, can be energy-intensive and expensive to operate, which concerns institutions with limited resources.
- **Legal and Ethical Issues:** The immutability of blockchain raises questions about compliance with data privacy laws, particularly the right to be forgotten.

6. Scalability Considerations for Blockchain in Libraries

For blockchain to be effectively scalable for digital preservation in libraries, a hybrid approach of **on-chain metadata storage** and **off-chain content storage** may be the most

viable solution. Blockchain can store hashes and metadata to ensure authenticity, while larger digital objects are stored in traditional repositories or distributed file systems such as IPFS (InterPlanetary File System).

7. Scope for Further Research

- Explore **cost-efficient blockchain models** (e.g., Proof of Stake) that consume less energy.
- Develop **interoperable blockchain frameworks** that integrate with existing digital preservation systems.
- Address **legal concerns** regarding blockchain's immutability in relation to privacy rights.

8. Conclusion

Blockchain has the potential to transform digital preservation in libraries by ensuring the authenticity, integrity, and provenance of digital records. However, its adoption at scale faces challenges related to cost, energy use, and storage limitations. Future research should focus on improving the scalability of blockchain systems for libraries, exploring hybrid models, and addressing legal and ethical considerations.

By integrating blockchain with other emerging technologies, such as distributed storage systems, libraries can leverage the best of both worlds to ensure the long-term preservation of digital content. Additional pilot projects and case studies will be crucial in testing these solutions in real-world library environments.

References

- Archangel: Securing our National Archives with AI and blockchain | University of Surrey. (n.d.). Retrieved October 14, 2024, from <https://www.surrey.ac.uk/news/archangel-securing-our-national-archives-ai-and-blockchain>
- Blockcerts. (n.d.). Blockchain Credentials. Blockcerts. Retrieved October 14, 2024, from <http://blockcerts.org/>
- Introduction: Blockchain Project – Blockchains for the Information Profession. (n.d.). Retrieved October 14, 2024, from <https://ischoolblogs.sjsu.edu/blockchains/project/>
- Jha, S. K. (2023). Application of blockchain technology in libraries and information centers services. *Library Hi Tech News*. <https://doi.org/10.1108/LHTN-02-2023-0020>
- Panda, S., & Kaur, N. (2023). Blockchain: A New Technology in Library System and Management. In B. Holland (Ed.), *Advances in Library and Information Science* (pp. 211–230). IGI Global. <https://doi.org/10.4018/978-1-6684-7693-2.ch011>
- Trček, D. (2022). Cultural heritage preservation by using blockchain technologies. *Heritage Science*, 10 (1), 6. <https://doi.org/10.1186/s40494-021-00643-9>

- Vivien Oluchi, E., Efemini, M., Oseni Yahaya, D., & Oladokun, B. D. (2023). Application of blockchain technology to 21st century library services: Benefits and best practices. *Data and Metadata*, 2, 59. <https://doi.org/10.56294/dm202359>
- Wayback Machine. (2024). In *Wikipedia*. Retrieved October 14, 2024, from https://en.wikipedia.org/w/index.php?title=Wayback_Machine&oldid=1251093688
- Yaga, D., Mell, P., Roby, N., & Scarfone, K. (2018). Blockchain technology overview (NIST IR 8202; p. NIST IR 8202). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.IR.8202>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

