PRESERVING THE PAST, ENSURING THE FUTURE: CHALLENGES AND INNOVATIONS IN DIGITAL PRESERVATION OF INFORMATION

NJUGUNA A., KIMOTE Z.
PhD Candidates- Kirinyaga University, KENYA
Correspondence : *andykim7119@gmail.com*

Abstract

The task of digital preservation, which entails safeguarding enormous amounts of knowledge for future generations, is crucial for contemporary society. The intricate field of digital preservation is examined in this paper, with a particular emphasis on how the variety of formats and sources calls for fundamental adjustments in preservation methods. In the ever evolving technological landscape, reactive and flexible methods are needed; traditional ones are no longer adequate. Concerning digital archiving, the primary objectives of this research are data integrity, obsolescence, and the ongoing threat of technological revolution. An extensive guide to long-term digital data storage is to be provided, along with an explanation of new methods and technologies created to address these concerns. An in-depth and comprehensive literature review and analysis of pertinent case studies form part of the methodology. Important subjects comprise metadata standards, emulation, format migration, and cooperative initiatives in the preservation community. Anticipated results include a thorough grasp of the challenges encountered by digital preservation initiatives as well as an investigation of effective strategies and resources to get past these roadblocks. This study enhances to the current of knowledge on digital preservation by pointing out obstacles and providing workable fixes for long-term preservation. It emphasizes the necessity of modifying conservation methods to take advantage of evolving technological capabilities and advocates for an active, future-focused approach to safeguarding the enduring legacy of our digital heritage.

Keywords: Digital Preservation, Format Migration, Data Integrity, Obsolescence, Technological Change

1.0 Introduction

The digital age has revolutionized the way information is created, disseminated, and preserved [1]. Worldwide, the creation, access, and dissemination of information have been completely transformed by the advent of digital technologies. The increase in digital information has thus created previously unheard-of opportunities for accessibility and dissemination, ranging from digital content and multimedia artifacts to digitized manuscripts and archival documents. However, with the boundless opportunities presented by digitalization, there exist formidable obstacles that endanger the enduring viability of digital assets.

Houghton [2] observed that organizational challenges in digital preservation include technical, administrative, legal, and logistical aspects. According to Tripathi [3] digital preservation requires collective measures incorporating traditional and digital means, including bundling, storage, quality control, and risk preparedness, to ensure the enduring preservation of digital objects. Elsewhere, Burda and Teuteberg [4] argued that digital preservation faces organizational challenges, such as lack of cost-benefit analysis and decision-making methods, and suggested a need for a research agenda for future exploration. This paper looks at the innovations and difficulties experienced in digital preservation with a particular emphasis on the value of saving historical documents, scientific knowledge, and cultural heritage for future generations.

Objectives

The objectives of this study were to analyze the challenges inherent in digital preservation, as identified in the existing literature, identify innovative strategies and technologies to address these challenges, makes recommendations for advancing digital preservation and highlight areas for future research and collaboration.

Methodology

This study uses a literature review of peer-reviewed papers, reports, and institutional documents on the subject, a thorough search of academic databases, institutional repositories, and pertinent websites. To narrow down the search results, terms like "format migration," "digital preservation," "long-term access," and "technological obsolescence" were used. Subsequently, a thematic analysis of the chosen literature was conducted to pinpoint recurring issues, creative solutions, and suggestions for improving digital preservation techniques.

Challenges in Digital Preservation

Preservation of digital information has become an important concern for institutions, organizations, and society at large in the rapidly changing digital age. While observing that digital assets must be protected for future generations he observed that digital preservation brings with it several complicated issues with organizational, legal, ethical and technological components [2]. This study examined the various obstacles that digital preservation initiatives must overcome and addresses the complexities of format migration, technology obsolescence, data authenticity, and sustainability and policy frameworks. Through a thorough analysis of these issues, there is a possibility of gaining a deeper understanding of the intricacies of digital archiving and identify tactics and solutions to effectively address these issues.

Technological Obsolescence

Technological obsolescence poses a significant challenge to long-term digital preservation. As observed by El Idrissi [5] long-term digital preservation faces challenges from software and format obsolescence, requiring fundamental preservation strategies to guarantee accessibility, authenticity, and integrity of digital objects for a long time. The author observed that digital materials are prone to format obsolescence and hardware/software dependencies due to their inherent fragility and susceptibility to rapid technological changes. Thus, the swift advancement of hardware and software

platforms presents a noteworthy peril to the sustained accessibility of digital assets, given that obsolete file formats and outdated technologies grow progressively incompatible with contemporary systems. Elsewhere, Bekele et al [6] cautioned that the integrity of our scientific knowledge and cultural heritage is at risk when digital materials become unreadable and inaccessible over time due to technological obsolescence if preventive measures are not taken [6].

Format Migration

Considering the wide variety of file formats used to encode digital content, format migration is another crucial issue in digital preservation. Xie et al. [7] proposed that there may be need to migrate digital assets to more sustainable formats to ensure their long-term viability as file formats become outdated or unsupported by evolving technology. He indicated that format migration and preservation metadata are crucial for the digital preservation of scientific data, enabling reproducible experiments and promoting research and innovation [7]. Data loss, corruption, and fidelity loss are just a few of the possible hazards associated with format migration, which is a difficult and resource-intensive process. Because of their intrinsic complexity and interdependencies, complex digital objects like multimedia files and interactive websites also pose unique preservation challenges.

Data Authenticity

One of the main issues with digital preservation is preserving the integrity and authenticity of digital materials. Thus, digital assets, in contrast to physical artifacts, are subject to unauthorized changes, manipulation, and tampering, which raises questions about the data's authenticity and reliability. Robust cryptographic checksums, audit trails to identify and stop unwanted modifications, and digital signature verification techniques are necessary to ensure the integrity of digital materials. Furthermore, proving provenance and authenticity metadata establishment is essential for recording the chain of custody and confirming legitimacy of digital materials over time.

Sustainability and Funding

According to Burda and Teuteberg [4] scarcity of resources and conflicting priorities, make digital preservation projects face formidable challenges in terms of funding and sustainability. Important investments in staff, software, hardware, and training are needed to create and maintain a sustainable digital archiving infrastructure [4]. Further, ongoing financial and other support from governmental bodies, charitable institutions, and other stakeholders is necessary to guarantee the long-term viability of digital repositories and archives. There are concerns about sustainability and longevity of digital preservation initiatives due to funding challenges.

Legal and Ethical Considerations

Digital preservation is heavily influenced by ethical and legal factors, which include things like copyright, privacy, and intellectual property rights. Corrado [8] observed that digital preservation faces challenges beyond just technology, including intellectual property and economic concerns, which impact software preservation and web archiving. The use and distribution of digital materials are governed by a complex legal environment that includes copyright restrictions, licensing agreements, and privacy regulations. These regulations present challenges for digital preservation initiatives. In addition, moral questions about the digitization and sharing of indigenous knowledge, culturally sensitive materials, and personal data present complex moral challenges that call for careful thought and discussion with the appropriate parties.

Innovations in Digital Preservation

Emulation

Replicating the original software environment in which digital materials were created and accessed is called emulation, and it is a promising method of digital preservation. Arunkumar and Devendran [9] suggested that a combination of migration and emulation can effectively preserve digital data, reducing dependence on future technology and ensuring its longevity. Thus, faithful reproduction and rendering of digital content in its

original context is made possible by emulation, which ensures a lasting accessibility and usability of digital assets by simulating outdated hardware and software platforms. Digital materials can thus be accessed and interpreted for eternity thanks to emulation, which provides a workable solution to the problems of format obsolescence and technological dependency.

Blockchain Technology

It has been demonstrated in a study how Blockchain-based data preservation system effectively preserves medical data, ensuring privacy and security while ensuring reliable treatment progress [10]. The authors indicated that decentralized and impenetrable mechanisms provided by blockchain technology for storage and verification of digital assets offer enormous potential for digital preservation. This technology can be used by digital preservation projects to create unchangeable records of ownership, provenance, and authenticity for digital materials. In another report, Bernal Barnabe et al. [11] observed that privacy-preserving solutions in blockchains are evolving to enable users to become anonymous and take charge of their private data during digital transactions, addressing challenges in diverse scenarios like eGovernment, eHealth, and Cooperative. This possibility increases accountability and trust in the preservation process. Additionally, strong protections against data loss, censorship, and unauthorized changes are provided by blockchain-based decentralized storage systems, guaranteeing the longevity and integrity of digital assets.

Artificial Intelligence

With automated methods for content analysis, metadata extraction, and preservation planning, Artificial Intelligence (AI) has the potential to completely transform digital preservation. In a report by Russo and Ciaccio [12] AI technologies, specifically Vision Transformer and Diffusion models, have the potential to enhance environmental monitoring and preservation by improving accuracy, efficiency, and communication. The authors observed that Artificial Intelligence (AI) driven algorithms are able to evaluate

enormous amounts of digital material, spot preservation hazards, and rank the importance and value of different materials in order of preservation. Furthermore, digital materials can be automatically transcribed, translated, and indexed with the help of AI-driven methods like machine learning and natural language processing, which will improve their discoverability and accessibility.

Decentralized Storage Solutions

Muthalibu [13] argued that digital asset preservation is made possible by decentralized storage solutions like distributed ledger technologies and peer-to-peer networks, which provide robust and censorship-resistant platforms in contrast to centralized storage systems that are susceptible to censorship and single points of failure, decentralized storage solutions disperse digital content among multiple nodes in a network, guaranteeing fault tolerance and redundancy. The resilience and longevity of digital assets can be improved by digital preservation initiatives by utilizing decentralized storage solutions to reduce the risks of data loss, censorship, and unauthorized alterations.

Persistent Identifiers

As distinct and long-lasting references to digital assets, persistent identifiers, or PIDs, are crucial tools in the field of digital preservation. Through consistent citation, expedited discovery, and easy access across various platforms and repositories, these identifiers play a vital role in guaranteeing the longevity of digital publications. PIDs are assigned to digital assets so that long-term access and management of these resources can be managed efficiently for the benefit of future generations under digital preservation initiatives. Di Cosmo et al. [14] pointed out that, by offering a trustworthy framework for locating and selecting digital assets, PIDs essentially play a major part in digital preservation since they make it easier to organize, cite, and retrieve digital content. Digital preservation efforts can thus ensure that digital content remains relevant and useful in an ever-changing digital landscape by putting in place a strong infrastructure

for its ongoing management and accessibility through methodical implementation of PIDs.

5. Case Studies

Digital information preservation is a difficult task in the rapidly changing digital world. Kapeliushna et al. [15] observed that globally, enterprises are facing the challenges of safeguarding digital resources, guaranteeing their availability, and preserving their authenticity in the long run. Prominent initiatives aimed at tackling these issues head-on include the Digital Preservation Strategy of the European Commission, the Digital Stewardship Alliance (NDSA), and the Digital Preservation Coalition (DPC). These groups lead the charge in ensuring sustainability and accessibility of digital materials for upcoming generations via cooperation, advocacy, and strategic planning.

The National Digital Stewardship Alliance- (NDSA)

Gallinger et al.[16] observed that, the goal of the NDSA is to advance digital stewardship in the US through a cooperative network of institutions and organizations focused on creation of standards, guidelines, and resources for digital preservation through cooperative projects, working groups, and advocacy initiatives. The Levels of Digital Preservation framework, one of the main projects of the NDSA, offers a tiered method for evaluating and enhancing digital preservation practices. The five levels of the framework, which go from basic to advanced, each correspond to a set of skills and tasks required for efficient digital preservation. Institutions can thus prioritize areas for improvement and evaluate their preparedness for digital preservation by implementing these levels of Digital Preservation framework.

Bringing together a broad community of professionals, including librarians, archivists, technologists, and researchers, the NDSA encourages cooperation and knowledge-sharing among its members through working groups, task forces, and annual meetings, which helps to develop best practices, standards, and resources for digital preservation. Additionally, the NDSA has further been instrumental in promoting national funding

and policy initiatives for digital preservation. The NDSA has developed capacity within the cultural heritage sector and increased awareness of the value of digital preservation through cooperative projects and partnerships.

6. Future Trends

Future growth and innovation in the field of digital preservation are anticipated due to the need to handle the ever-changing opportunities and challenges of the digital age. Emerging trends and fresh approaches will reshape the field of digital preservation in the years to come, propelled by the innovative work of institutions devoted to safeguarding our digital legacy. Georgopoulos [17] observed that the future of digital preservation is expected to be shaped by transformative technologies and cooperative efforts aimed at safeguarding the lasting accessibility and relevance of the digital cultural heritage. These efforts will range from adoption of decentralized and blockchain-based systems to improve the integrity and longevity of digital assets, to advancements in artificial intelligence and machine learning for automated metadata curation and enrichment.

a) ***Integration of Emerging Technologies*****:** Incorporation of new technologies, including Artificial Intelligence (AI), Machine Learning and Blockchain, is one of the most significant future paths in digital preservation. Woods and Lee [18] noted that these technologies have a great deal of promise for streamlining metadata extraction processes, keeping digital materials authentic and intact, and automating archiving workflows. Utilizing AI-driven algorithms, businesses can increase the effectiveness and scalability of their digital preservation initiatives by optimizing digitization, metadata generation, and content analysis processes.

b) ***Enhanced Collaboration and Interoperability*****:** Digital preservation will continue to advance primarily due to collaboration and interoperability as observed by Hilbert [1]. Collaboration between heritage institutions, research organizations, and technology providers to share resources, expertise, and best practices will be crucial as the volume and complexity of digital materials continue to rise.

Furthermore, promoting standards-based methodologies and interoperability will make it easier to exchange and access digital materials across various platforms and repositories.

c) ***Focus on Diversity, Equity, and Inclusion*:** Initiatives for digital preservation will place more of a focus on inclusion, equity, and diversity in the coming years. Organizations will ensure that digital collections reflect a wide range of cultural, linguistic, and geographical diversity because they understand how important it is to preserve diverse voices and perspectives. Additionally, addressing inclusivity and accessibility concerns will be a top priority in order to guarantee that digital materials are usable by all users, including those with limited technology access or disabilities.

d) ***Sustainable Preservation Strategies*:** Organizations working to create sustainable preservation methods that balance social, economic, and environmental factors will continue to prioritize sustainability in their digital preservation efforts. The importance of reducing the environmental impact of digital preservation activities—such as through energy-efficient storage options and ethical e-waste management—will only increase. Furthermore, the long-term viability of digital preservation initiatives will be enhanced by programs that enable communities to take charge of their digital legacy and develop local competence.

e) ***Ethical and Legal Considerations*:** Boughman in a study pointed out that digital preservation techniques will continue to be shaped by ethical and legal issues, especially in relation to copyright, privacy, and intellectual property rights [18]. To maintain appropriate management of digital artifacts, organizations will need to negotiate complex legal frameworks and ethical standards as indicated by Corrado in a study about legal considerations on a digitization undertaking. Furthermore, it is imperative that digital preservation programs prioritize addressing cultural sensitivity, indigenous rights, and community permission [8]. This will necessitate stakeholder consultation and ethical standards' observance.

f) ***Collaborative Partnerships*:** The future of digital preservation initiatives will be greatly influenced by collaborative collaborations [20]. Thus, the need for collaboration across institutions, industries, and disciplines is growing as the complexity of digital preservation concerns rises. In order to maximize the pooling of resources, technology, and knowledge, future paths in collaborative partnerships will require cultivating cross-sectoral cooperation involving government agencies, academic researchers, corporate stakeholders, and cultural heritage institutions. Proposals such as the DPC-Digital Preservation Coalition and the NDSA- National Digital Stewardship Alliance serve as excellent examples of collaborative approaches that work well for exchanging knowledge, developing capacity, and creating best practices for digital preservation.

g) ***Capacity Building and Training*:** Initiatives aimed at increasing capacity and providing training will remain crucial in providing experts with the abilities and information required to handle the challenges associated with digital preservation. Expanding access to education and training programs is a future direction for capacity building and training, especially in underprivileged areas and communities. Professionals can learn about new trends and technologies, improve their digital preservation skills, and share best practices with colleagues through programs like webinars, online courses, and workshops. Stakeholders can guarantee the sustainability of digital preservation efforts and empower a new generation of professionals in the field by funding capacity building and training initiatives.

h) ***Policy Development and Advocacy*:** Legal, regulatory, and funding frameworks pertaining to digital preservation will be significantly shaped by policy development and advocacy as pointed out by Corrado [8]. He suggested that advocating for policies that support open access, data sharing, and long-term sustainability in digital preservation initiatives is one of the future directions in policy development and advocacy [8]. Advocacy groups and policymakers can be

crucial in promoting the adoption of best practices and standards, advocating for funding to support preservation initiatives, and educating the public about the value of digital preservation. Digital preservation advocates can impact policy decisions and guarantee that digital preservation stays a top priority on national and international agendas by interacting with legislators, advocacy groups, and stakeholders. Professionals in preservation and guarantee the long-term viability of digital preservation initiatives.

7. Findings

Digital preservation poses complex challenges, including technological obsolescence, data integrity, and sustainability, as identified in numerous studies. Innovative strategies and technologies proposed in the literature, such as emulation, blockchain technology, and decentralized storage, offer promising solutions for addressing these challenges and ensuring the long-term preservation of digital resources. Collaboration among libraries, archives, museums, and technology developers is consistently emphasized in the literature as essential for advancing the field of digital preservation and developing interoperable standards and tools.

Conclusion

The literature assessment cited in this study demonstrate the vital relevance of digital preservation in conserving cultural heritage and intellectual materials for future generations. By adopting new trends and technologies presented in the literature, institutions and practitioners can overcome the obstacles of digital preservation and assurelong-term accessibility and usability of digital resources. Moving forward, coordinated efforts and persistent funding are required to advance digital preservation while effectively preserving communal memory and knowledge. Digital preservation is thus, a complicated and varied enterprise that requires ongoing collaboration, creativity, and investment. As the rapidly changing digital landscape is navigated, it becomes

increasingly important to address the numerous issues associated with safeguarding digital information.

References

1. Hilbert M. Digital technology and social change: the digital transformation of society from a historical perspective. Dialogues Clin Neurosci. 2020;22(2):189–94.

2. Houghton B. Preservation challenges in the digital age. DLib Mag. 2016;22(7/8).

3. Tripathi S. Digital preservation: some underlying issues for long-term preservation. Libr Hi Tech News. 2018;35(2):8–12.

4. Burda D, Teuteberg F. Sustaining accessibility of information through digital preservation: A literature review. J Inf Sci. 2013;39(4):442–58.

5. El Idrissi B. Long-term digital preservation: A preliminary study on software and format obsolescence. In: Proceedings of the ArabWIC 6th Annual International Conference Research Track. New York, NY, USA: ACM; 2019.

6. Bekele MK, Pierdicca R, Frontoni E, Malinverni ES, Gain J. A survey of augmented, virtual, and mixed reality for Cultural Heritage. J Comput Cult Herit. 2018;11(2):1–36.

7. Xie J, Zhang M, Ma Y. Using format migration and preservation metadata to support digital preservation of scientific data. In: 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS). IEEE; 2019.

8. Corrado EM. Digital preservation is not just a technology problem. Tech Serv Q. 2022;39(2):143–51.

9. Arunkumar K, Devendran A. Digital Data Preservation—A Viable Solution. In: Data Management, Analytics and Innovation. Singapore: Springer Singapore; 2019. p. 129–41.

10. Li H, Zhu L, Shen M, Gao F, Tao X, Liu S. Blockchain-based data preservation system for medical data. J Med Syst. 2018;42(8).

11. Bernal Bernabe J, Canovas JL, Hernandez-Ramos JL, Torres Moreno R, Skarmeta A. Privacy-preserving solutions for blockchain: Review and challenges. IEEE Access. 2019;7:164908–40.

12. Russo P, Ciaccio FD. Recent advances in AI for enhanced environmental monitoring and preservation. In: 2023 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters (MetroSea). IEEE; 2023.

13. Muthalibu A. Decentralised Code Sharing Platform. International Journal of Scientific Research in Engineering and Management. 2022;06(08).

14. Di Cosmo R, Gruenpeter M, Zacchiroli S. 204.4 Identifiers for Digital Objects: The case of software source code preservation. Open Science Framework; 2022.

15. Kapeliushna T, Goloborodko A, Nesterenko S, Bezhenar I, Matviichuk B. Analysis of digitalization changes and their impact on enterprise security management under uncertainty. Nauk Visnyk Natsionalnoho Hirny Hoho Uniwersytetu. 2023;(4):150–6.

16. Gallinger M, Bailey J, Cariani K, Owens T, Altman M. Trends in digital preservation capacity and practice: Results from the 2nd bi-annual national digital stewardship alliance storage survey. DLib Mag. 2017;23(7/8).

17. Georgopoulos A. Contemporary digital technologies at the service of cultural heritage. In: Heritage Preservation. Singapore: Springer Singapore; 2018. p. 1–20.

18. Woods K, Lee CA. Acquisition and processing of disk images to further archival goals. Archiving. 2012;9(1):147–52.

19. Baughman MS. Issue on the transformation of scholarly communications. Res Libr Issue. 2015;(287):1–1.

20. Demetres MR, Delgado D, Wright DN. The impact of institutional repositories: a systematic review. J Med Libr Assoc. 2020;108(2).