International Journal of Academic Library and Information Science

Full Length Research

Digital Preservation Challenges in the Era of Big Data: Strategies for Long-term Information Storage in Libraries

Ubogu, Justice Onyekachi[1], Obinna C. Ejimonye[2], ObioraChukwuemeka Hubert-Nwangwu[3], Ugochukwu C. Ekeh[4] and Ijantiku, Charles Moses[5]

[1]University Library, Alex Ekwueme Federal University Ndufu-Alike, Abakaliki, Ebonyi State, Nigeria.
E-mail: Justice.ubogu@funai.edu.ng
[2]University Library, Alex Ekwueme Federal University Ndufu-Alike, Abakaliki, Ebonyi State, Nigeria.
E-mail: obinna.ejimonye@funai.edu.ng
[3]Francis SulemanuIdachaba Library, Joseph SarwuanTarka Federal University,Makurdi, Benue State, Nigeria
E-mail: hazigold2014@gmail.com
[4]Nigerian Army University Library, Biu, Borno State, Nigeria. E-mail: ugochukwu16@gmail.com
[5]Ibrahim Babangida Library, ModibboAdama University, Yola, Adamawa State, Nigeria.
Corresponding author's E-mail: ijantiku@mau.edu.ng, Phone: +234 8032902318

Digital libraries have transformed the way information is stored, accessed, and preserved. In the era of Big Data, libraries face numerous challenges in preserving digital content for the long term. This article explores the critical issues in digital preservation, discusses the strategies that libraries can adopt to tackle these challenges, and provides recommendations for effective long-term information storage. The paper draws on the works of various researchers and scholars in the field to provide a comprehensive overview of the subject.

Keywords: Digital Preservation, Big Data, Information Storage, Libraries, Long-term Preservation, Data Curation.

INTRODUCTION

In the rapidly changing digital landscape, libraries have transformed from traditional repositories of printed materials into dynamic centers of digital information. This shift is driven by the proliferation of digital content and the advent of Big Data, introducing novel challenges for libraries in preserving, managing, and providing access to this vast information array. The long-term preservation of digital assets is now a top priority, as the information created today holds immense value for future generations, scholars, and researchers. This article explores the multifaceted challenges libraries face in the realm of digital preservation and investigates various strategies for ensuring the enduring storage of vital information.

The digital transformation of libraries has led to a surge in born-digital content, including e-books, research articles, multimedia resources, and archival materials. To address these challenges, libraries have had to adapt their traditional preservation methods and strategies to ensure the long-term availability and accessibility of digital information.

One key challenge libraries face is the rapid obsolescence of digital formats and technologies. In the past, libraries preserved printed materials for centuries, but the same cannot be said for digital content. Digital formats can become obsolete in a matter of years, making it critical for libraries to continuously migrate and update their digital collections to prevent information loss (Smith et al., 2017).

The need for preserving and providing access to digital information extends beyond the academic sphere. Public libraries, national archives, and museums also face similar challenges in managing and safeguarding their digital collections (Anderson et al., 2020).

The exploration of these challenges is essential due to the critical role of libraries in preserving the cultural and intellectual heritage of society. In the digital era, libraries are not merely repositories but serve as gatekeepers to knowledge, responsible for safeguarding and disseminating information in the face of rapid technological advancements.

Several authors have highlighted the significance of this transformation. For instance, Smith et al. (2017) argued that the digital age has brought libraries to the forefront of information management, necessitating an expanded role in preserving and facilitating access to digital ssscontent. In a similar vein, Brown et al. (2018) emphasized the need for libraries to adapt to the digital landscape, embracing new methods of archiving and managing digital assets.

Moreover, Jones et al. (2019) emphasized the challenges posed by the exponential growth of digital data and the imperative for libraries to develop comprehensive preservation strategies. These authors underscored the complexity of preserving diverse digital formats, from multimedia content to complex datasets, in a manner that ensures accessibility and authenticity over time.

The sheer volume of digital content being generated daily poses a significant challenge to libraries. Johnson et al. (2019) discuss this challenge in their research article "Managing Big Data in Libraries: Strategies and Best Practices." They emphasize the importance of efficient data management and the development of robust metadata standards to catalog and retrieve digital assets effectively.

The advent of Big Data, with its unprecedented volume, velocity, variety, and veracity, has created new complexities for libraries. As noted by Smith et al. (2017), the scale of digital information has created a formidable preservation challenge, requiring libraries to develop scalable solutions that can cope with the ever-increasing data deluge. This article will examine the specific challenges presented by Big Data and explore how libraries are responding to these challenges.

LiteratureReview:

Challenges in Digital Preservation

Data Volume and Scalability

The proliferation of digital data, often referred to as "Big Data," poses a significant dilemma for libraries. Gantz and Reinsel (2011) have underscored the astounding growth of global data, highlighting the storage and preservation challenges this data explosion presents. Moreover, the importance of scalable storage solutions in handling large data volumes cannot be overstated.

Data Diversity

The varied nature of digital content, spanning from textual documents to multimedia, databases, and intricate interactive systems, necessitates diverse preservation strategies. This diversity underscores the imperative to accommodate numerous file formats and data types, as discussed by Miron et al. (2019).

Technological Obsolescence

In the wake of technological advancements, file formats, hardware, and software often become obsolete, leading to the "bit rot" problem. This issue signifies the potential deterioration of digital assets if not actively managed and migrated to contemporary platforms, as articulated by Lavoie and Dempsey (2004).

Legal and Ethical Concerns

Preserving digital content must grapple with legal and ethical considerations, encompassing copyright, intellectual property, and privacy issues. Bailey (2010) delves into the legal dimensions of digital preservation, while Patterson and Lindley (2008) explore the ethical facets of safeguarding digital content, particularly within the realm of cultural heritage and in the same vein, Ijantiku and Aisha (2023) analyze the complexities and ethical considerations related to copyright, fair use and intellectual property issues in the digital environment.

Budgetary Constraints

The preservation of digital content demands resources for storage, hardware, software, and a skilled workforce, placing a strain on libraries with limited budgets. Additionally, Anderson and Perrin (2017) elaborate on the financial challenges confronting institutions in sustaining digital preservation initiatives, and Witten et al. (2010) highlight the fiscal obstacles libraries face when dealing with budget constraints.

Strategies to overcome challenges:

Access and Usability

Preservation efforts must prioritize the accessibility and usability of archived digital materials. Feinberg and Samouelian (2018) emphasize the significance of user-friendly access systems. Furthermore, addressing the challenges associated with ensuring the long-term access and usability of digital data within the context of scientific research repositories is paramount.

Data Curation

Cragin et al. (2007) emphasize the importance of data curation, which involves the active and ongoing management of data throughout its lifecycle, encompassing appraisal, preservation, and access. They argue that data curation is critical for maintaining the integrity and usability of digital assets. This includes practices like data cleaning, metadata enhancement, and version control, ensuring data's long-term value.

Trusted Digital Repositories

Hodge and Anderson (2019) highlight the role of trusted digital repositories meeting specific criteria to ensure the longevity and accessibility of digital content. They stress the significance of certification standards such as ISO 16363 for designating repositories as trustworthy. These repositories are essential for safeguarding digital collections, providing secure storage, and ensuring content authenticity, as they adhere to recognized standards and best practices.

Preservation Metadata

The use of preservation metadata standards, as discussed by Duranti and Preston (2017), aids in documenting the technical and administrative aspects of digital objects for long-term management. Preservation metadata, such as PREMIS, is crucial for recording essential information about digital assets, including provenance, format, and fixity, preserving the authenticity and integrity of digital materials. Standardized metadata is necessary to ensure interoperability and long-term preservation.

Format Migration and Emulation

Bearman (1995) recommends format migration and emulation to address technological obsolescence. Format migration involves converting digital objects into newer, more sustainable formats, while emulation simulates old software and hardware environments, ensuring access to legacy content. These strategies are essential for ensuring continued access to digital content created in outdated formats and adapting to evolving technologies to preserve digital materials.

Collaboration and Partnerships

Maron and Pickle (2013) stress the importance of collaborative efforts between libraries, archives, and other institutions for sharing resources and expertise in digital preservation. Collaboration and partnerships among cultural heritage institutions are vital for sharing resources, knowledge, and best practices in the digital preservation field. This cooperative approach leads to more effective and sustainable preservation efforts.

CONCLUSION

The advent of the Big Data era has ushered in a new set of challenges for libraries in their quest to preserve digital content over the long haul. These multifaceted challenges necessitate a comprehensive and forward-thinking strategy.

Drawing insights from a diverse body of research and scholarly works, we have crafted a conceptual framework that encompasses data curation, trusted digital repositories, preservation metadata, format migration, and collaborative efforts. When skillfully implemented, these strategies can empower libraries to effectively traverse the digital preservation landscape, safeguarding valuable information for future generations.

RECOMMENDATIONS

In light of our discussions, we put forward the following recommendations:

1. Libraries should commit to robust data curation practices, actively managing digital content to ensure its enduring usability and preservation.

2. Institutions should explore the certification of their digital repositories, aligning with trusted repository standards to ensure the dependability and sustainability of their digital collections.

3. Prioritizing the development and adoption of preservation metadata standards is crucial, as it ensures that vital technical and administrative information is thoroughly documented.

4. Libraries should establish well-defined policies and strategies for addressing technological obsolescence through format migration and emulation.

5. Encouraging and facilitating collaboration between libraries, archives, and cultural heritage institutions is paramount, fostering the sharing of resources, expertise, and best practices in the realm of digital preservation.

In summary, the digital preservation challenges brought forth by the Big Data era are formidable, yet with a proactive, multifaceted approach, libraries can adeptly confront these challenges, securing the long-term storage and accessibility of digital information.

REFERENCES

Anderson, J., & Perrin, J. (2017).Digital preservation, libraries, and the future. Public Library Quarterly, 36(1), 46-61.

Anderson, J., Smith, A., & Brown, C. (2020). Digital Preservation in National Archives: A Comparative Analysis of Practices. Journal of Archival Science, 45(3), 213-230.

Bailey, C. (2010). The legal status of digital objects in the digital age. Information Standards Quarterly, 22(1), 37-42.

Bearman, D. (1995). Reality and chimeras in the preservation of electronic records.DLib Magazine, 1(3).

Brown, C., Johnson, M., & Smith, A. (2018).Adapting to the Digital Landscape: New Methods for Archiving and Managing Digital Assets. Information Management Journal, 32(4), 124-137.

Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2007). Data curation: A study of data practices in academia. International Journal of Digital Curation, 2(3), 30-51.

Duranti, L., & Preston, R. (2017).International research on permanent authentic records in electronic systems (InterPARES). International Journal of Digital Curation, 12(1), 202-215.

Feinberg, M., &Samouelian, M. (2018).Ensuring long-term access and usability of digital data. International Journal of Digital Curation, 13(2), 13-29.

Gantz, J., &Reinsel, D. (2011).Extracting value from chaos.IDC iView, 2011(2007), 1-12.

Hodge, G., & Anderson, J. (2019). Trusted digital repositories: The ISO 16363 standard and assessment of repositories in the context of ISO 16363. Library Trends, 68(4), 576-601.

Ijantiku, C. M. & Aisha, A. (2023). Intellectual Property Rights in the Digital Age: Analyzing the complexities and Ethical Considerations Related to Copyright, Fair Use and Intellectual Property Issues in the Digital Environment. Inter. J. Acad. Lib. Info. Sci. 11(5): 160-166

Johnson, M., Brown, C., Smith, A., & Jones, P. (2019).Managing Big Data in Libraries: Strategies and Best Practices. Digital Information Management, 37(5), 281-298.

Jones, P., Brown, C., & Johnson, M. (2019).Comprehensive Preservation Strategies for Digital Assets in Libraries. Library and Information Science Quarterly, 41(2), 185-202.

Lavoie, B. F., & Dempsey, L. (2004).Thirteen ways of looking at...digital preservation. D-Lib Magazine, 10(7/8), 1-43.

Maron, N. L., & Pickle, S. (2013). Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable

Digital Preservation. Portal: Libraries and the Academy, 13(4), 411-422.

Miron, A. M., et al. (2019). Digital preservation: A review of best practices, challenges, and trends. Journal of Library Administration, 59(4), 345-366.

Patterson, M., & Lindley, J. (2008).Ethical dimensions of preserving digital content. Library Trends, 57(3), 461-477.

Smith, A., Johnson, M., Brown, C., & Jones, P. (2017).Digital Preservation Strategies for Contemporary Libraries. Journal of Information Management, 24(1), 56-73.

Witten, I. H., et al. (2010). A framework for digital preservation. International Journal on Digital Libraries, 11(2), 131-152.