Debunking the Myth of Obsolescence: Strategies for Digital Heritage Conservation

Li Chuqiu

MA Culture and Heritage Management, University of Lincoln, Brayford Pool, Lincoln LN6 7TS, United Kingdom

28828002@students.lincoln.ac.uk

Abstract. This essay extensively explores the multifaceted challenges and complexities associated with preserving digital cultural heritage, using the BBC Domesday Project as a case study. It underscores the critical importance of meticulous planning, future-proofing, and proactive strategies in digital archival endeavors. Emphasizing the significance of adopting open standards and establishing clear data ownership provisions, the essay aims to mitigate risks of obsolescence and safeguard digital cultural heritage for future generations. Furthermore, it addresses the concept of obsolescence in digital preservation and proposes strategic recommendations to prevent potential catastrophic issues, ensuring sustained access and relevance over time. In addition, the document delves into the criteria for selecting materials for digitization, highlighting the importance of metadata infrastructure in digital archiving. It discusses the challenges associated with constructing digital cultural repositories and emphasizes the necessity of balancing various priorities, including cultural significance, rarity, fragility, research potential, accessibility, user-friendliness, and institutional objectives during the selection process. Moreover, the essay underscores the pivotal role of metadata in ensuring the longevity and accessibility of digital resources and archives. Overall, this essay serves as a comprehensive guide for cultural heritage institutions, offering valuable insights and recommendations for the preservation and interpretation of digital heritage objects.

Keywords: digital preservation, metadata, Dublin Core, EAD, MODS, digitizing heritage, authenticity, obsolescence

1. Introduction

Launched in 1984, the BBC Domesday Project serves as an instructive case study in digital heritage conservation. Domesday earned its moniker, "the day of judgement," for its immutable decisions, akin to those rendered on the day of judgement itself [1]. it was created for the 900th anniversary of the Domesday Book, consisted of two interactive video discs showcasing approximately one million individuals across the UK, along with video clips, images, and maps. Published by BBC Enterprises in November 1986, it was a significant interactive project, comprising the Community Disc and the National Disc. The accessioned material includes analogue image data converted into DigiBeta format and the original 12-inch video discs [2].

BBC Doomsday Project served as a notable instance of format drift and hardware obsolescence is exemplified [3] the software required to read the laserdiscs had become outdated technology [4]. Contrasting with the near-millennial longevity of its namesake, the Domesday Book, this digital counterpart succumbed to obsolescence in under a decade [5]. In addition, employing a proprietary format and hardware, resulted in accessibility and preservation challenges inherent to closed systems [6]. Subsequently, the BBC Domesday Reloaded initiative, launched in 2011, endeavored to enhance accessibility by digitizing the content and offering it online, albeit still reliant on proprietary technology [7]. An open approach could have fostered increased community engagement and ensured long-term preservation by enabling user contributions and sustainability efforts beyond the BBC's scope [8]. By leveraging open standards and platforms, such as those seen in open-source initiatives, the project could have facilitated broader participation and innovation, ultimately enhancing its accessibility and resilience over time [9].

This analysis dissects the deficiencies of the project and provides strategic suggestions to avoid similar pitfalls in subsequent digital archival endeavors. Adopting a cross-disciplinary approach, this investigation endeavors to untangle the complexities involved in safeguarding digital cultural heritage. And also aims to examine, from an archival standpoint, the factors that should

be taken into account prior to endorsing any digitization proposal, along with the challenges that may arise during the implementation process.

2. Challenges in Constructing Digital Cultural Repositories

Digitalizing heritage presents distinct benefits. They enable direct information and content delivery to end-users, facilitating remote retrieval. Moreover, the quality of images can be notably high, often augmented, with ongoing advancements enhancing capabilities [10]. However, digital archive initiatives frequently encounter a spectrum of challenges, spanning from financial limitations and technological hurdles to legal and ethical considerations, as well as unique organizational contextual factors [11]. Also, one significant challenge of digitizing heritage is the preservation of fragile or non-standard physical artifacts [12]. Chief among all the challenges is technological obsolescence.

The Domesday disc is anticipated to encounter limitations, with a primary concern being its "read-only" nature, precluding easy updates. There was contemplation regarding a revised edition aimed at rectifying significant omissions and errors. However, the feasibility of implementing a comprehensive and regular update process remains uncertain, contingent upon the response elicited by the initial edition [13]. The concept of the "wiki" model, inspired by collaborative platforms like Wikipedia, offers a potential solution to this issue by enabling continuous updates and revisions by users, Wikis are dynamic websites that allow for full editing capabilities, enabling users to freely access, read, restructure, and modify both textual and visual content, including text and images [14].

Digital media, unlike their physical counterparts, face the inherent risk of technological obsolescence due to the fast-paced evolution of the digital landscape. This vulnerability stems from their reliance on specific hardware and software configurations, which undergo constant advancements and iterations. Thus, another significant challenge to the long-term preservation of digital information is also the issue of "rapidly changing storage devices" [15]. As a result, the longevity and accessibility of digital media are inextricably linked to the sustained compatibility and interoperability of these underlying technological components.

The process of digitizing cultural assets involves a careful balance of procedural and technical deliberations. Thus choosing the most suitable digitization technique (scanning, digital photography, or three-dimensional modeling) is contingent on the nature of the items being digitized. Archivists should not only comprehend the ever-evolving software and hardware landscape but also with acquiring knowledge in museology and related disciplines to navigate the complexities of their roles effectively [16]. Furthermore, integrating the viewpoints and requirements of end-users into the digitization process is also vital to ensure that the resulting digital assets align with their needs and expectations [17].

Furthermore, the safeguarding of digital artifacts' authenticity and integrity is critical. A genuine archive constitutes a contextually grounded organic repository of evidence, rather than a compilation of disparate information. Hilary Jenkinson, a prominent figure in archival theory, emphasized the significance of authenticity to archives, defining it as the principle that archives are "preserved in official custody . . . and free from suspicion of having been tampered with."[18]. The absence of robust validation mechanisms can result in digital repositories that are vulnerable to corruption or unauthorized alterations, compromising their authenticity and the trust placed in them. With the increased reliance on digital resources, individuals and institutions face the challenge of verifying the authenticity of digital images or reproductions. Given the ease of electronic manipulation, ensuring the integrity of such materials has become an exceedingly urgent concern [19].

3. Metadata Infrastructure for Digital Archiving

Metadata is the linchpin of digital collections, enabling efficient location, access, and contextual understanding of digital items. Besser [15] suggests that metadata serves as the primary safeguard against obsolescence and degradation of digital information. Institutions can mitigate future inaccessibility risks by providing comprehensive metadata, capturing essential technical, descriptive, and administrative details. This framework aids in long-term management, interpretation, and access of digital objects. A well-conceived metadata strategy captures critical descriptive, administrative, and structural data. Embracing standardized metadata frameworks like Dublin Core, EAD and MODS bolsters interoperability and eases integration with sophisticated CMS platforms. These systems must be versatile, supporting a range of media types and equipped with features for version control and access management to effectively oversee and disseminate digital heritage content.

The concept underlying Dublin Core aimed to streamline resource discovery. Specifically, it was devised as a system to enhance Internet users' ability to locate web pages that closely aligned with their interests, surpassing the precision of search engines available at the time of its inception [20]. The other two prominent metadata schemas widely used for describing digital material in cultural heritage exhibit differentiation and redundancy. Encoded Archival Description (EAD) is a differentiated schema tailored specifically for archival description, while Metadata Object Description Schema (MODS) offers a more generalized approach serving as an intermediary between simple and complex formats [21].

EAD stands as the most widely adopted metadata schema for archival description [22]. As a differentiated standard, it is designed for encoding archival finding aids with Extensible Markup Language (XML) [23], enabling precise and optimized accessibility to archival resources across various institutions for users. In contrast, MODS introduces an element of productive redundancy by serving as a supplementary metadata format, offering an intermediary option between simplistic formats like Dublin Core, which have minimal fields and lack substructure, and intricate formats like MARC 21, which feature numerous data elements

with diverse structural complexities [24]. One study also proposes an EAD to MODS crosswalk, leveraging further redundancy by involving the semantic mapping of elements between these two metadata standards. It aims to better facilitate the archival description and management of digital resources, providing a bridge between the specialized archival metadata format of EAD and the more flexible and interoperable MODS schema [21]. The differentiation enables each schema to incorporate specialized elements and structures, enhancing the significance and utility of metadata within its specific domain [25].

Metadata plays a pivotal role in ensuring the longevity and enduring accessibility of digital resources and archives [26]. Terry Cook underscores the importance of retaining fundamental archival principles such as evidence, context, and provenance amidst technological advancements. Rather than discarding these principles, we should acknowledge how technology transforms them [27]. By embracing best practices in metadata creation, standardization, and integration with sophisticated digital preservation systems, cultural heritage institutions can safeguard their digital holdings, ensuring that these invaluable resources remain accessible, authentic, and contextually rich for generations to come.

The BBC Domesday Project's failure underscores the importance of meticulous planning and future-proofing in digital archival endeavors. Several factors contributed to its downfall, including reliance on proprietary technologies, inadequate documentation, and a lack of long-term preservation foresight. Legal complexities surrounding proprietary formats and ownership restrictions also pose significant barriers to sustained access and preservation efforts [28]. Addressing these challenges may involve emulating legacy hardware and software environments within modern systems, albeit requiring negotiation of legal permissions or development of workarounds for proprietary technologies [29]. Alternatively, data migration from obsolete to newer formats is often feasible (Moss & Currall, 2004), but requires careful consideration of potential intellectual property barriers or licensing terms associated with the original formats [30]. These setbacks advocate for a proactive approach in future endeavors, emphasizing open standards and clear data ownership provisions to mitigate obsolescence risks and safeguard digital cultural heritage for posterity.

4. Roadmap for Future Digital Archival Initiatives

The process of digital archival unveil the multifaceted contextualization of a digital historical representation. Each connection or potential association among units of historical information, curated through a selection process, search queries, and archival provenance, embodies an act of historical interpretation by the representation's creator. This confluence signifies the merging of historiographical and archival decisions [31].

Based on the analysis of BBC Domesday, this report proposes the following roadmap for future digital preservation projects and provides advice on how to avoid potential catastrophic issues, which include: the decision to digitize entails considerations regarding the rationale behind digitization, the selection criteria for materials to be digitized, the target audience, strategies for packaging and discovery of digitized content, and ensuring the sustainability of the digitization initiative as a whole:

4.1. Rationale for Digitization

The digitization of cultural heritage and the subsequent digital interactions surrounding these resources cannot be viewed as a neutral or apolitical endeavor, detached from the complex power dynamics and historical contexts that shape our societies. As Pickover [32] aptly highlights, this process is inherently embedded within intricate networks of power relations, ideologies, and socio-cultural influences. It is crucial to recognize that the motivations and objectives driving digitization efforts are not merely technical or logistical considerations but are profoundly shaped by larger historical, political, and cultural forces. The decision to digitize certain materials while potentially overlooking others, the framing and representation of these resources, and the mechanisms for enabling access and engagement are all influenced by underlying power structures, cultural biases, and dominant narratives. Consequently, a clear and critical understanding of the rationales and motivations behind digitization initiatives is crucial. These rationales may range from preserving cultural heritage and safeguarding historical narratives to enabling broader access to materials and facilitating research and analysis.

4.2. Selection Criteria

Determining which materials should be prioritized for digitization constitutes a critical juncture in the preservation process. Strategic selection methods form a fundamental aspect of digitization projects, guided by considerations such as the nature and intellectual content of collections, their condition and usage patterns, as well as the copyright status of original materials, all in alignment with institutional strategy (Hazen et al. 1998). Notably, the selection process for digital conversion emphasizes accessibility and user-friendliness over physical preservation concerns. As Gertz [33] asserts, projects should be guided by specific objectives rather than being solely driven by technological capabilities. Institutions must strategically select collections for digitization based on predefined objectives, subjecting each decision to a rigorous cost-benefit analysis to evaluate its potential impact on both end-users and the institution itself [34].

This multifaceted approach to selection underscores the necessity of balancing competing priorities, such as cultural significance, rarity, fragility, research potential, accessibility, user-friendliness, and institutional objectives. By carefully weighing

these factors and conducting thorough cost-benefit analyses, institutions can optimize their resource allocation and maximize the impact of their digitization initiatives, thereby safeguarding invaluable cultural heritage resources while fostering greater public engagement and scholarly inquiry.

4.3. Target Audience

Identifying the intended audience for the digitized content is essential for tailoring the presentation, discoverability, and accessibility of the materials. Different user groups, such as researchers, students, or the general public, may have varying needs and expectations. Smith [35] also highlights the importance of promoting social inclusion within heritage organizations to involve the public in engaging with authoritative versions of historical and archaeological narratives, which often reflect the interests guided by expert-driven principles.

4.4. Content Packaging and Discovery

As a cultural heritage archivist, it's imperative to implement effective strategies for organizing, describing, and presenting digitized content to ensure its discoverability and usability. While the Dublin Core standard represents progress toward interoperability, it's essential to maintain consistency in entering attribute contents across all digitized resources. This consistency enhances the effectiveness of search functionalities and improves user experience [36]. Given the exponential growth of the internet, relying solely on text searching may not meet users' needs, especially when search engines struggle to differentiate between objects with similar names. To address this challenge, cataloging using keywords and controlled vocabularies becomes crucial, enabling clearer distinction and more accurate retrieval of relevant resources [37].

Furthermore, the design of metadata standards and user interfaces should prioritize user-friendly navigation and access. By implementing these strategies, cultural heritage institutions can optimize the discoverability and usability of digitized content, thereby enriching the user experience and maximizing the impact of their digital collections.

4.5. Sustainability

Ensuring the long-term sustainability of the digitization initiative is paramount. This involves considerations related to data management, storage infrastructure, rights management, and ongoing funding and resource allocation.The majority of contemporary digital storage media possess relatively brief lifespans. Recordable CDs and DVDs typically last between 5 to 10 years, while hard disks have a lifespan of approximately 20 to 30 years [38]. While this situation is expected to improve over time, the current reality necessitates the meticulous and periodic verification and rewriting of the bitstream for every digital object onto new media. This imperative arises from the inherent vulnerability of digital data to degradation and corruption, a phenomenon colloquially termed "bit rot." [39]. To mitigate the risks associated with bit rot and ensure the enduring integrity of digital artifacts, a multi-pronged approach is essential. First, it is crucial to maintain multiple independent copies of each digital object, preferably on diverse types of storage media [37]. This redundancy serves as a safeguard, enabling the recovery and restoration of corrupted or compromised data by cross-referencing against the remaining intact copies.

Furthermore, the regular and systematic monitoring and verification of the bitstream's integrity are paramount. This process involves employing robust data integrity checks and validation mechanisms to detect any deviations or anomalies in the digital bitstream. Upon the identification of such issues, immediate action must be taken to migrate the affected digital objects to new storage media, thereby preserving their fidelity and preventing irreversible data loss.

5. Conclusion

In the contemporary digital landscape, the preservation and interpretation of heritage objects have gained heightened significance due to the escalating risk of their loss stemming from technological advancements. Hence, the scrutiny of the BBC Domesday Project and subsequent analysis of strategies for conserving digital heritage have become imperative, this report analyzed the multifaceted challenges and intricacies involved in preserving digital cultural heritage. The Domesday Project's failures highlight the importance of proactive strategy, standards compliance, and stakeholder participation in digital preservation. Future initiatives must prioritize meticulous planning, digitization techniques, and safeguarding artifact authenticity. Metadata, particularly standardized frameworks like Dublin Core, EAD, and MODS, is crucial for ensuring the longevity and accessibility of digital archives.

The concept of obsolescence in digital preservation has been challenged, and strategic suggestions have been proposed to circumvent potential catastrophic issues, ensuring enduring access and relevance for future generations. By embracing best practices in metadata creation, standardization, and integration with sophisticated digital preservation systems, cultural heritage institutions can safeguard their digital holdings, ensuring that these invaluable resources remain accessible, authentic, and contextually rich for posterity. however, It is important to note that there is no singular correct method; methodologies may vary [40]. The lessons learned from the BBC Domesday Project serve as a sobering reminder of the necessity for a proactive and

considered approach in future digital archival endeavors, aiming to circumvent obsolescence and safeguard digital cultural heritage for generations to come.

References

[1]   Wood, M. (1999). *Domesday*. Ebury Press. Retrieved from http://books.google.ie/books?id=TW1nAAAAMAAJ&q=Domesday+A+Search+for+the+Roots+of+England&dq=Domesday+A+Search+for+the+Roots+of+England&hl=&cd=2&source=gbs_api

[2]   Data from BBC Domesday Project | The National Archives. (2009, August 12). Retrieved from https://discovery.nationalarchives.gov.uk/details/r/C16160

[3]   Hessler, J. C. (2007, June 14). Archive Fever. *Journal of Map and Geography Libraries.* https://doi.org/10.1300/j230v03n02_05

[4]   Skeem, D. (2006). Digital Preservation in a Techno World: Issues Involved and Institutions Actively Participating. *Preservation.*

[5]   Finney, A. [no date]. BBC Domesday Project. Retrieved from https://www.atsf.co.uk/dottext/domesday.html.

[6]   BBC Domesday Project - Source code available - stardot.org.uk. (2017, October 14). Retrieved from https://stardot.org.uk/forums/viewtopic.php?t=13883

[7]   Data from BBC Domesday Project | The National Archives. (2009, August 12). Retrieved from https://discovery.nationalarchives.gov.uk/details/r/C16160

[8]   Balestrini, M., Bird, J., Marshall, P., Zaro, A., & Rogers, Y. (2014, April 26). Understanding sustained community engagement. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* https://doi.org/10.1145/2556288.2557323

[9]   Bhatt, P., Ahmad, A. J., & Roomi, M. A. (2016, June). Social innovation with open source software: User engagement and development challenges in India. *Technovation, 52–53,* 28–39. https://doi.org/10.1016/j.technovation.2016.01.004

[10] Conway, P. (2000). Overview: Rationale for digitization and preservation. In M. Sitts (Ed.), *Handbook for digital projects: A management tool for preservation and access.* Retrieved February 20, 2001 from the World Wide Web: http://www.nedcc.org/digital/dighome.htm

[11] Mckay. (2003). Digitization in an Archival Environment. *Electronic Journal of Academic and Special Librarianship (Winter 2003), 4*(1).

[12] Pandey, R., & Kumar, V. (2020, April 1). Exploring the Impediments to Digitization and Digital Preservation of Cultural Heritage Resources: A Selective Review. *Preservation, Digital Technology & Culture, 49*(1), 26–37. https://doi.org/10.1515/pdtc-2020-0006

[13] Goddard, J., & Armstrong, P. (1986). The 1986 Domesday Project. *Transactions of the Institute of British Geographers, 11*(3), 290. Available at: http://dx.doi.org/10.2307/621789.

[14] Husain, Noushad & Introduction,. (2010). Wiki: A Web Based Technology for Reflective, Collaborative and Constructive Learning in Higher Education. *University News, AIU.*

[15] Besser, H. (2001). Digital preservation of moving image material. *The Moving Image, 1*(2), 39-55.

[16] Hoopes, J. (1997). The future of the past: Archaeology and anthropology on the World Wide Web. *Archives and Museum Informatics, 11*(2), 87-105.

[17] Schmidt, J., Drews, P., & Schirmer, I. (2016). End-users' perspective on digitalization: A study on work order processing in the German banking industry.

[18] Jenkinson, Hilary. (1965). *A Manual of Archive Administration.* London: Percy Lund, Humphries & Co. Ltd.

[19] Beamsley, T. (1999). Securing digital image assets in museums and libraries: A risk management approach. *Library Trends, 48*(2), 358-78.

[20] Beall, J. (2004), "Dublin Core: An Obituary", *Library Hi Tech News, 21*(8), 40-41. https://doi.org/10.1108/07419050410567399

[21] Bountouri, L., & Gergatsoulis, M. (2009, August 31). Interoperability Between Archival and Bibliographic Metadata: An EAD to MODS Crosswalk. *Journal of Library Metadata, 9*(1–2), 98–133. https://doi.org/10.1080/19386380903095107

[22] Library of Congress (LC). (2002a). EAD Version 2002 official site. Retrieved December 13, 2008, from http://www.loc.gov/ead/index.html

[23] W3C. (2006). Extensible Markup Language (XML) 1.0 (4th ed.). Retrieved December 13, 2008, from http://www.w3.org/TR/xml/

[24] Guenther, R. S. (2003, January). MODS: The Metadata Object Description Schema. *Portal: Libraries and the Academy, 3*(1), 137–150. https://doi.org/10.1353/pla.2003.0006

[25] Eichler, R., Giebler, C., Gröger, C., Hoos, E., Schwarz, H., & Mitschang, B. (2021, July 2). Enterprise-Wide Metadata Management. *Business Information Systems,* 269–279. https://doi.org/10.52825/bis.v1i.47

[26] Sugimoto, S. (2014, January 2). Digital archives and metadata as critical infrastructure to keep community memory safe for the future – lessons from Japanese activities. *Archives and Manuscripts, 42*(1), 61–72. https://doi.org/10.1080/01576895.2014.893833

[27] Cook, T. (1997). What is Past is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift. *Archivaria, 43,* 17-63. Retrieved from https://archivaria.ca/index.php/archivaria/article/view/12175

[28] Hirtle, P. B. (2003). Digital preservation and copyright.

[29] Byers, F. R. (2003). Care and Handling of CDs and DVDs—A Guide for Librarians and Archivists. *NIST Special Publication, 500*(252), 1-40.

[30] Steenbakkers, J. F. (2005). Digital archiving in the twenty-first century: Practice at the national library of the Netherlands. *Library Trends, 54*(1), 33-56.

[31] Sternfeld, J. (2011, September 1). Archival Theory and Digital Historiography: Selection, Search, and Metadata as Archival Processes for Assessing Historical Contextualization. *The American Archivist, 74*(2), 544–575. https://doi.org/10.17723/aarc.74.2.644851p6gmg432h0

[32] Pickover, Michele. (2014). "Patrimony, Power and Politics: Selecting, Constructing and Preserving Digital Heritage Content in South Africa and Africa". Paper presented at the IFLA WLIC – Libraries, Citizens, Societies: Confluence for Knowledge Conference, Lyon, France, August 16–22.

[33] Gertz, J. (2000). Selection for preservation in the digital age: An overview. *Library Resources & Technical Services, 44*(2), 97-104.

[34]   de Stefano, P. (2000). Selection for digital conversion. In A. Kenney & O. Rieger (Eds.), *Moving theory into practice: digital imaging for libraries and archives* (pp. 11-23). Mountain View, CA: Research Libraries Group.

[35]   Smith, Laurajane. (2009). "Class, Heritage and the Negotiation of Place." Paper presented at the Missing out on Heritage: Socio-Economic Status and Heritage Participation Conference for English Heritage, London, March 23.

[36]   DC-dot. (n.d.). https://www.ukoln.ac.uk/metadata/dcdot/

[37]   Moss, M., & Currall, J. (2004). Digitisation: Taking stock1. *Journal of the Society of Archivists, 25*(2), 123–137. https://doi.org/10.1080/0037981042000271457

[38]   Rosenthal, D.S.H. (2010), "Format obsolescence: assessing the threat and the defenses", *Library Hi Tech, 28*(2), 195-210. https://doi.org/10.1108/07378831011047613

[39]   Cerf, V. G. (2011, June). Avoiding "Bit Rot": Long-Term Preservation of Digital Information [Point of View]. *Proceedings of the IEEE, 99*(6), 915–916. https://doi.org/10.1109/jproc.2011.2124190

[40]   Kenney, A., & Rieger, O. (2000). Introduction: Moving theory into practice. In A. Kenney & O. Rieger (Eds.), *Moving theory into practice: digital imaging for libraries and archives.* Mountain View, CA: Research Libraries Group.