DOI: https://doi.org/10.52783/tojqi.v11i1.9969

Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 11, Issue 1, January 2020: 369-376

Preserving For Posterity- Use of Digitization Techniques for Libraries

Anupriya
Department of School of Computing, Graphic Era Hill University, Dehradun, Uttarakhand, India 248002

Abstract

The content of decaying physical relics has been digitised as a preservation approach. At the same time, it opens up new possibilities for knowledge exchange and exploitation by reducing the constraints imposed by physical confinement. How we produce, organise, and disseminate knowledge has been revolutionised by the widespread use of digital formats. However, it is not a simple task to guarantee the longevity of digital media. Valuable data generated in the digital age may be lost forever if a satisfactory solution is not found. In this article, we discuss some of the challenges inherent in the process of digitising and archiving digital material. We outline the economic condition that provides the backdrop for further efforts in this direction.

Keywords: digital library, digital archive, metadata, standards, XML, preservation.

Introduction

As information and communication technologies have spread, there has been a rapid movement away from archiving material in analogue formats like paper manuscripts, pictures, and audio-visual recordings and towards digital ones. The proliferation and convenience with which we can create, distribute, and share digital material has revolutionised the way we process data. We understand the value of timely, accurate data and are committed to creating tools that will allow us to maximise our use of it. We use several methods of digitization to convert paper documents into digital files that may be used in conjunction with information that was created using a digital device from the start.

"Digitization is the act of transforming information material from a conventional format into a digitally readable version," as stated in Harrod's Librarians Glossary (Peytherch, 2000). Digital imaging encompasses not just the scanning of printed books but also the capture of other types of artwork such as manuscripts, paintings, prints, etc. Because fragile originals do not need to be handled after digitization, it has become increasingly popular as part of the push to improve preservation techniques. As digitised texts can be easily searched and digital images of dissipated or difficult the originals can be machine enhanced, access has been significantly enhanced.

Several major libraries across the country have initiated digitization initiatives aimed at fragile books, fragile materials, and rare books.

The goal of digital preservation, as defined by Russell (2009), is to maintain the accessibility, longevity, and authenticity of digital material for future generations. A more exact definition is "the

storage, maintenance, and accessibility of a digital asset across time, often as a result of implementing a number of digital preservation solutions, which may include tech preservation, technological emulation, or data transfer."

Digital Archiving – Preserving the Posterity

Social, financial, cultural, and intellectual legacy data are archived digitally in what are known as "digital archives" for future use. The process of digital archiving involves locating and migrating dormant data from the live production system or databases to more permanent storage. Archives save digital or born-digital records for posterity because of their potential evidentiary, legal, contextual, cultural, historical, or informative significance. The data is saved so that it may be used in the future. Digital archiving makes it possible to keep information safe for the long term, make it accessible to anybody who needs it, encourage paperless offices, and protect sensitive information from prying eyes. The museums, libraries, and archives of the memory institutions save knowledge and cultural artefacts for future generations. The UN Educational, Scientific, and Cultural Organisation (UNESCO) has requested its member nations to record and keep a record of the effects of COVID-19. Records of government agencies, military units, traditions, patents, the past, and culture are preserved at the public archives of several countries, including Australia, New Zealand, the UK, Spain, Germany, Switzerland, the United States, India, and Canada. Their histories, traditions, and experiences are recorded and stored in communal archives for posterity (Rodrigues, 2016).

PDA, or personal digital archiving, is a relatively recent idea. It's a term used when either a person or an institution like a library or museum is interested in preserving an individual's digital data. However, depending on the information and resources at their disposal, different individuals and organisations may choose to implement varying methods of controlling and safeguarding digital archives. Personal digital assistants (PDAs) combine the best of digital archiving and data management (Condron, 2019).

Archival documents may be safeguarded against many types of damage, both natural and manmade, thanks to preservation and conservation efforts. They are an essential tool for preserving and disseminating human knowledge for the current and future generations. Archives store records of historical significance, which are vital to the development of a flourishing society. Archives have a civic duty to protect and keep safe the records of past achievements in the arts, sciences, and other fields for the benefit of future generations. Rare books, manuscripts, and artworks increase in value with time, and it's next to impossible to replicate them exactly. Therefore, it is crucial that archival records be preserved and conserved.

Issues of Digitization in conversation and preservation

Archive and library administrators, curators, technical personnel, and the specific users who are accountable for the digitization process confront a number of challenges and must be mindful of their actions at all times. The problem may be seen in both positive and negative light. Mass digitalization has not altered the core services provided by archives and libraries, but the administration of digital assets brings new obstacles in terms of employing and maintaining complicated gear and software. They have a lot of problems and worries, including:

1) Concerning the Law,

2) Lack of proper training and human error,

3) Format Shifts and the Decline of Old Technologies,

4) Money,

5) Renewing and harmonising,

6) Continually Evolving Tech, Which Includes But Is Not Limited To:

7) Simulation

A complete infrastructure for safeguarding and preserving digital assets has various technological, organisational, legal, and economic problems despite indications of increased concern about digital preservation.

(i) Technology problems: The effectiveness of digital preservation relies on the use of suitable technological medium. However, the hardware and software technologies are evolving rapidly, rendering the current technology fragile and obsolete. In order for the system to provide the desired results, the relevant institution will need to update the necessary hardware and software.

(ii) Nature of Content: There is a wide variety of digital data formats. Materials that can't be duplicated in conventional hard-copy, such as interactive web pages, maps, and so on, may present a challenge, but the majority of items that are true replicas of their print document, such as books, reports, correspondence, etc., can be transformed into electronic formats and preserved.

(iii) Machine Dependency: Due to the machine-dependent nature of digital material, it may be necessary to utilise the same computer and software that was utilised to create the information in order to access it. However, components of preservation systems should be updated and revised as computer and storage technologies evolve. The continuous availability of digital material depends on this, and its absence might pose the biggest technical danger.

(iv) Sustainability of Digital Object : Since digital content is typically stored on magnetic and optical media, which quickly degrade, it is especially susceptible to loss and destruction. In addition, it is vulnerable to unexpected breakdowns brought on by heat, humidity, air pollutants, or a malfunctioning reading or writing apparatus. When compared to traditional, hard-backed copies, digital objects sometimes have a very limited lifespan. This might be an issue for archiving digital materials.

(v) Inappropriate budget: To preserve digital media, an institution has to invest in expensive new equipment and hire specially educated staff. In this instance, the success of the digital archive project may be jeopardised if sufficient funding is not allocated. Intellectual Property Libraries have a fantastic potential to enhance their services thanks to the digital technologies. It also enables novel strategies for the long-term storage and distribution of library holdings. Intellectual Property Rights, however, may impede both the transformation of originality and the dissemination of the same, since digitalization often entails the movement of unique items. A lack of qualified and technically savvy personnel is impeding efforts to preserve digital materials. However, it can be challenging to find suitable personnel at times. There is no legislative framework in the country for the upkeep and protection of resources of national significance, which is problem number eight on the list. Another

obstacle to successful digital preservation is the general lack of both understanding and resources in this area. The inherent properties of digital storage media—whether magnetic or optical—lead to a decline in quality that is more fast than that of printed materials. Digital data is machine-dependent, therefore certain software and hardware components of a computer are required for correct "reading." The devices and software we formerly relied on suddenly become outdated and useless. The most common issues with digital preservation are medium failure or degradation and the fast advancements in computing technology that often render legacy systems ineffective.

Preserving Valuable documents and artifact

This is why it comes as no surprise that content such as books, records, and films have been digitised. Digitization methods unshackle information from its physical containers. The increased information agility afforded by digital formats has far-reaching practical, political, legal, and economic consequences for society. The preservation of information from decaying paper artefacts, film archives, and sound recordings has led to the widespread use of digitalization as a method. OCR, ASR, V&I processing, and allied fields have all seen increased activity as a result of this process's technological needs. Concurrently, the limits of conventional IT architectures are being tested by the demands of managing the digitised data, which in turn is dramatically altering pricing models. Digitization comes at a hefty price. Digitising and applying OCR to a single newspaper page may cost as much as one pound. About 750 million pages, some as old as the 18th century, make up the British Library's regional, national, and international newspaper collection. We need to find economic drivers that may drive the digitalization of ageing data in order to make this endeavour cost-effective.

Given that the government likely won't be able to foot the entire bill for digitising content, what's in it for private companies to get involved? Information that is relatively new is likely to be the most valuable to companies, but much of this data is protected by copyright and must be handled carefully. It's obvious that solving a number of difficult problems is crucial to the project's success. Below, we discuss how the increasingly cutthroat online search industry and the advertising sector connected to it have impacted recent developments in content digitization.

Over the past decade, digitization efforts have been undertaken by nearly every major library. Typically, government organisations or private donors underwrite such endeavours. It's intriguing that funding from Web search companies have recently boosted their digitising initiatives. Revenue for online information services comes from users clicking on adverts on the sites they see after doing a search. The most popular subjects and information on the Internet account for just a tiny fraction of search queries, yet those few inquiries account for the vast majority of search service income (80%). The primary commercial goal is, therefore, to increase the company's proportion of Internet searches. Increasing user loyalty by providing a better online experience is a common tactic for accomplishing this goal. Online services are improving the range and quality of the material they offer in search results, in addition to users interface changes and branding made possible by the browser extensions. They are working with publishers to enable access to newly produced premium content, as opposed to just gathering and indexing publicly accessible Web data. Libraries and archives are looking for new methods to digitise and upload historical materials online.

Investment in Content Digitization

There has been a frenzy of content digitization efforts because to the increased competitiveness in the Web search business. Google signed agreements in the summer of 2005 to scan and index select materials from the libraries at Stanford, Harvard, and Michigan as well as the Public Library of New York and the Oxford College Library in the United Kingdom. The process of scanning results in the production of two copies, both of which may be indexed and searched using Google Books. Copyright issues necessitate limiting content choices to public domain or otherwise non-protected works. Scanning facilities have been set up specifically for this purpose. On October 3, 2005, "the Internet Archive, Yahoo! Inc., Adobe Systems Inc., the European Archive, HP Labs, the National Archives (UK), O'Reilly Media Inc., Prelinger Archives, the University of California at Berkeley, and the University of Toronto came together to form the Open Content Alliance (OCA)" (http://opencontentalliance.org), a global consortium dedicated to preserving the rights of content creators while allowing for open access to that content. Digital print and multimedia materials from libraries, archives, and publishers relating to cultural, historical, and technical topics is among the OCA's primary responsibilities.

MSN joined an open Content Alliance in October 2006, and the two organisations have since collaborated to digitise 100,000 volumes from the British Library. 2.1.2 European Union Effort at Coordination On April 28, 2005, six EU Member States asked the EU Commission to organise an effort to harmonise and organise national digitization activities throughout Europe in response to the first evidence of commercial attempts to interact with university libraries to invest in digitization.

On September 30, 2005, in response, the European Union Commission issued a Communication paper and a request for online public consultation titled "i2010: Digital Libraries," seeking the public's opinion on crucial matters related to the electronic preservation of national heritage. The communication document quotes statistics showing that in 2001 there were 2,533,893,879 books as well as bound periodicals (volumes) within European libraries (EU 25), which serves to illustrate the scale of the challenge faced by EU countries in preserving the artistic and patrimony of European nations. The audiovisual materials are of particular importance since the analogue formats degrade with time and result in content loss. There is a broad variety of proposals in the consultation answers, representing the many diverse stakeholders' interests and points of view on digitalization and preservation. The reader is referred to the complete report for more information. In this section, we discuss the problems that, in our opinion, are at the crux of the content digitalization difficulty. The massive scale of the digitization project necessitates a dedicated and methodical approach to the key problems of longevity, accessibility, and preservation. We anticipate some level of coordination can aid the effort, such as the creation of reporting and information services that hold data on all the content that has already been digitised, or at least about the material that incurs high the digitization cost so that duplication of effort ought to be avoided wherever possible. To advance the digitalization process in a sustainable and economically viable way, we think it is crucial to create a robust ecosystem centred on the information services sector. The sheer volume and interconnectedness of the variables at play. Physical artefacts' state and rate of degradation, digitization costs, content relevancy, and legal, social, and technical barriers to access are all factors to consider. Information's worth is relative to the circumstances in which it is presented, especially in light of recent developments and pressing concerns. As a result, historical data is most useful when combined with up-to-date information.

However, data associated with regular occurrences can be used to draw the right conclusion. This means that your selection process has to be in sync with your content exploitation models—specifically, the model that estimates the amount of interest in a certain field of study and determines how to maximise its utility by combining it with previously collected data. Combining historical data with instructional materials is one such instance. An obvious value-add is created when textbooks are supplemented with digital versions of relevant archival materials; this premium on educational resources may be used to fund further digitization efforts. Integrating and connecting to the educational curriculum is essential. In addition, at the national level, instruction in subjects like history, language, geography, or literature is often conducted in the original tongue with an emphasis on nationalistic elements of the collective past. Therefore, educational settings are ideally suited to the promotion of the digitalization of materials in the native language. Copyright & DRM (Digital Rights Management) are tools for protecting value distribution. Two components of technology, digital rights management (DRM) and micro-payment technology, are required to guarantee that publishers and writers may recoup the value of the work that has been published. They will be more willing to share data online if they have some say over the money being made from it. A reliable service will adhere to all applicable laws and regulations, safeguarding the rights of writers and publishers of copyrighted materials.

Preserving Digital Media

Too frequently, the long-term preservation of digital material takes a backseat to the immediate usage and administration of digital information. Unlike paper records, which will still be available 50, 20, or even 10 years from now with the right preservation plan, digital information from today will be inaccessible when that time period has passed. Ironically, the popularity of digital material might end up hurting it in the long run. New formats and storage medium are constantly being introduced as a result of technical advancements that make the creation and consumption of digital material easier. Old file types and software soon become obsolete. We risk losing access to digitally-created content if we don't use sufficient technology and best practises to make sure it can be read and used in modern information environments. This is a difficult problem with deep roots in the history of our culture. If we don't do anything about it, it will also weaken the method we've chosen to digitise paper and media materials in order to preserve knowledge from physical artefacts.

Problems and Initiatives

Bergman, Lyman, and Varian report that the annual value of EU-produced digital documents at risk of digital degradation is more than €3 billion. The financial impact on corporations and governments is enormous.

National museums and libraries, who have legal obligations to protect digital materials, have been particularly vocal in sharing knowledge about the present preservation landscape and inspiring new approaches to tool development and process design.

The problems of preservation are more than can be met by a single organisation. The European Union Commission recognises the importance of preserving cultural and scientific materials for future generations, and has allocated funds under the Sixth Foundation Programme to this end. The objective is to get archives, libraries, and research institutes to work together so that they may bring their unique viewpoints and expertise to bear on the issue.

Conclusion

Since most people in the 21st century would rather look up answers to their questions on the internet than rummage through filing cabinets full of paper documents, digitization has become increasingly important. The longevity of the information and data generated in the digital age depends on figuring out how to keep it in digital form for the foreseeable future.

References

1. Beatrice Ayodeji Fabunmi and Others (2009): Digitization of Library Resources: Challenges and Implications For Policy and Planning, International Journal of African & African American Studies Vol. V, No. 2, Jul 2006 p. n.2, https://www.researchgate.net/publication/228521970
2. Bergman, M.K.: Untapped Assets: The $3 Trillion Value of U.S. Enterprise Documents. BrightPlanet CorporationWhite Paper, July (2005) http://www.brightplanet.com/pdf/DocumentsValue.pdf .
3. Commission of the European Communities: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. i2010: Digital Libraries, Brussels, 30. Sept. (2005) http://europa.eu.int/information_society/activities/digital_libraries/doc/communication/en _comm_digital _libraries.pdf .
4. Commission of the European Communities: Communication Staff Working Document. Annex to the Communication from the Commission. i2010: Digital Libraries. Questions for online consultation. Brussels, 30. Sept. (2005) http://europa.eu.int /information_society/activities/digital_libraries/doc/communication/an nex2_en.pdf .
5. Commission of the European Communities: Results online consultation 'i2010: digital libraries', Brussels, 2. March (2006) http://europa.eu.int /information_society/activities/digital_libraries/doc/communication/re sults_of_online _consultation_en.pdf
6. CommitteeDraft 1, 22. March (2004) http://xml.coverpages.org/OpenOfficeSpecificationV10-CD.pdf
7. Condron, M. (2019). Identifying individual and institutional motivations in personal digital archiving. Preservation, Digital Technology & Culture, 48(1), 28- 37.
8. Darlington, J., Finney, A., Pearce, A.: Domesday Redux: The rescue of the BBC Doomsday Project videodiscs. Ariadne, Issue 36, July (2003) http://www.ariadne.ac.uk/issue36/tna/
9. Duval, E., Hodgins, W., Sutton, S., Weibel, S.L.: Metadata Principles and Practicalities. D-Lib Magazine, Vol. 8, Num. 4, April (2002) http://www.panjabdigilib.org
10. Lorie, R., van Diessen, R.: UVC: A Universal Computer for Long-Term Preservation of Digital Information. RJ 10338, IBM Almaden Research Center, San Jose, CA (2005)
11. Lyman, P., Varian, H. R.: How Much Information. The Journal of Electronic Publishing. December, 2000 Vol. 6, Issue 2, December (2003) http:// www.press.umich.edu/jep/06-02/ yman.html
12. McClelland, M., McArthur, D., Giersch, S., Geisler, G. Challenges for Service Providers when Importing Metadata in Digital Libraries. D-Lib Magazine, Vol. 8, Num. 4, April (2002) 10.
13. Mukhlesur Rahman and Muhammad Mezbah-ul-Islam (2012): Issues and Challenges for Sustainable Digital Preservation Practices in Bangladesh, International Seminar on 'Digital Libraries for Digital Nation'. 201 Organized by Library Association of Bangladesh, Date: October 17-18, 2012. Venue: Dhaka, Bangladesh Page 86-88.

14. Nur Atiqah Mohd Azim and Others (2018): Digitization of Records and Archives: Issues and Concerns, International Journal of Academic Research in Business and Social Sciences, Vol. 8, No. 9, p. n. 172. http://dx.doi.org/10.6007/IJARBSS/v8-i9/4582 15. Ibide. p. n.175-177. 16. Md.
15. Prytherch, R. J. (2000). Harrod's librarians' glossary and reference book: a directory of over 9,600 terms, organizations, projects, and acronyms in the areas of information management, library science, publishing, and archive management. (9th Ed.). Aldershot, Hants., England: Gower.
16. Raju, A.A.N. (2008) State of conservation and preservation of documents in public libraries. In: Ramaih, L.S.& Sujatha, G. (eds.). Preservation of library, archival and digital documents: Problems and perspective. New Delhi: Ess Ess Publication.
17. Rodrigues, A. (2016). Introducing an archival collecting model for the records created by South African Portuguese community organizations. Archives and Manuscripts, 44(3), 141-154. 13.
18. Russel, K. (n.d.). Digital preservation: Ensuring access to digital materials into the Future. CENDI cites resource collection. Retrieved May 12, 2009, romhttp://cendicites.infointl.com /item337.html