GLOBAL JOURNAL OF COMMERCE & MANAGEMENT PERSPECTIVE
(Published By: Global Institute for Research & Education)
www.gifre.org

The Case for Digital Preservation of the Press and Audio-Visual Documents in Mauritius

Nirmal Kumar BETCHOO
Université des Mascareignes, Pamplemousses, 00230, Mauritius.

Abstract

While the Mauritian Archives, under the aegis of the Ministry of Information and the National Library of Mauritius, have been constantly being challenged by the threat of stocking information in the form paper documents, it would be wise to consider how content technology in this particular sector that contributes to the cultural patrimony of the country should be digitally preserved. Alongside, the audio-visual sector should also be subject to digital preservation. This problem is current to emerging economies including Mauritius where information has been preserved using traditional methods over the decades where they have provided a wealth of information to researchers, students and decision makers. The threat of data loss is imminent regarding the archival of data in its primitive form. Concerning audio-visual data, such threat could be immense and deprive future generations of resources that are excellent in reviving past memories and major events. Mauritius has, so far, not been able to develop content technologies for the digital preservation of data. This research paper aims at identifying the simple methods of preserving data while at the same time considering how data could be digitally preserved and made customisable to the needs of potential users. Digital preservation would certainly bridge the gap between what the past and the future where the wealth of data preserved becomes a historical and cultural legacy.

Key words: *digital preservation, press and audio visual, Mauritius.*

1. Introduction

The preservation of press and audio-visual documents is a must for every society because it allows data to be preserved and used for the coming generations. Mauritius has had a fairly good system of traditional archiving of information if this concerns the written media. Under a traditional system of archiving newspapers, books and related publications, content preservation has been successful. This is not an end in itself because preservation of paper documents becomes very difficult over the decades through the degradation of pulp that forms the essential component of newspapers. There is also the importance of preserving audio-visual documents in Mauritius which are mainly stored in the Mauritius Broadcasting Corporation (MBC), so far, the only organisation that deals with archiving national information. Although more sectors are getting involved in audio-visual production, the MBC might still have an important role to play in data storage and preservation.

This paper aims at discussing the relevance of data preservation within two national organisations namely the National Library of Mauritius and the Mauritius Broadcasting Corporation. Regarding the audio-visual sector this could also be addressed to the ex-Mauritius College of the Air now known as the Open University of Mauritius.

The key objective is to consider the relevance of data preservation in a digital form and see how developments in digital technology would efficiently apply to the Mauritian context where the problem of storing information is already an important issue that needs to be well addressed because of the ongoing challenge of storing more information in today's knowledge-based society.

2. Literature Review

The literature review regarding digital preservation will focus on two areas separately in order to clearly define and explain the purpose of preserving data in a digital format in two areas where the technology of preservation varies significantly. The first part of the literature review analyses the written media sector.

The written press and preservation needs

It is important to understand at the outset the term "digital preservation" essentially means. Digital preservation is described as the "series of managed activities necessary to ensure continued access to digital materials for as long as necessary". Institutions seeking to preserve digital materials must understand that preservation requires planning, care, and coordination over time [1]. However, due to their inherent fragility, historic newspapers have long been subject to deterioration. Efforts to microfilm historic newspapers have helped to ensure their preservation, but microfilm is not as accessible as items digitally distributed over the Internet [2].

The newspaper is one of the few chronicles of a defining event in history - or in our personal lives - that we are able to hold in our hands. Newspapers in Mauritius remain so far the most important source of secondary and historical data. Librarians and historians preserve newspapers from around the world in carefully controlled environments and transfer them to microfilm or electronic images [3].

Efficient newspaper preservation implies understanding the meaning of preservation in general, as well as understanding specific approaches, methods and techniques that are included in newspaper preservation concepts

[4].Newspapers in Mauritius have so far been preserved in library stores with the risk of deterioration that has become inherent over the years due to generally poor storage and handling conditions. Modern newsprint contains lignin and wood pulp. It's inherently acidic and doomed to become brittle and discolour over time when left unprotected and exposed to heat, humidity, light or pests. Careful storage and a few common sense precautions can dramatically slow deterioration [5].

Preservation of Digital Audio Visual Content

The second part of this literature review analyses the importance of data preservation in the audio-visual area. This is considered as a vital sector in the information area because most of the key historical developments of Mauritius or any other country depends to the extent to which past data has been preserved. Digital preservation can be understood as the series of managed activities necessary to ensure continued access to digital information for as long as necessary[6], involving the planning, resource allocation, and application of preservation methods and technologies to ensure that digital information of continuing value remains accessible and usable [7] It combines policies, strategies and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change. The goal of digital preservation is the accurate rendering of authenticated content over time [8].

From a broad perspective, it is seen that animated or live broadcasting data is missing regarding the past. Some short films were made by independent amateurs while most of the animated data came from the nation broadcasting station, MBC, which still holds monopoly in the broadcasting of national news. With the partial liberalisation of the media in 2002[9], part of the control of media storage has been passed on to independent news companies which mainly benefit from digital technological support.

In the turn of the 20th century, all broadcasting archive centres thought they were confronted with the same questions touching the future of their heritage: fragility of support, changes of technical formats, obsolescence of the machines of reading, etc. Initially MBC television adopted the 8mm or Super 8 mm and later the 16mm colour reels before moving to digital Beta SP system for professional video broadcast. In the new millennium, national television shifted to digital television broadcast under DVB by early 2006. Besides, the digitisation of broadcasting and of production, the necessity to be able to exploit analog data precipitated the need for appropriate means of preservation of data [10]. This applied to the Mauritian context but there was greater concern whereby a significant part of past audio-visual information was destroyed. In this concern, it remains imperative for the national television and upcoming media to think about preservation.

Public television is responsible for the production, broad-cast and dissemination of programs which form the richest audio-visual source of cultural history [11]. By nature and necessity, public broadcasting is an assortment of media types and formats. In whatever manifestations these objects previously existed, they become bits and bytes before they reach the public eye [12]. That is an enormous amount of digital information to manage over time. As we move into the increasingly complex digital world, those charged with preserving our television heritage have the opportunity to develop and establish better coordinated and standardised preservation policies and practices to ensure what television programs and related assets survive [13].

3. Methodology

This research is mainly based from studies and research made from external sources. The digitalisation of information does not call for specific methodologies for research given that the solutions expected from the problem are tackled in the very much same way as external organisations do. There is apparently no new or tailor-made approach to specific countries because digitalisation of content is done in the same way.

The researcher has reviewed a selected number of companies or organisations that have opted for digitalisation of their media and, more importantly, the preservation of such data. The discussion is made in two sections; firstly the written media and secondly the audio-visual sector. This has been developed as such since it is impractical to find out a common solution to both sectors which might be interrelated but are independent seen from the archiving and preservation of data perspective.

Based from research and studies done from the external environment, the researcher was able to develop two distinctive models in order to synthesise the literature and general observations made. The formulae were made distinctive to each form of media: written and audio-visual in order to avoid confusion and keep the model as simple and easy to interpret and understand. The following section looks out for developments regarding digitisation of media from the external environment including authoritative organisations like the Library of Congress and several European broadcasting media.

4. Developments in the Written Media

Efficient newspaper preservation implies understanding the meaning of preservation in general, as well as understanding specific approaches, methods and techniques that are included in newspaper preservation concepts.

Preservation Microfilming

Microfilm has been a preferred medium for preservation for the last fifty years and is likely to remain so for the foreseeable future. When processed and stored in accordance with the agreed standards, silver halide microfilm is very long lasting. If the film should exhibit signs of deterioration, a duplicate can be made to replace it. It will also be possible to digitize a deteriorating film, enhance the image quality, and then produce a new film but there may be some loss of quality [14]. The single most critical factor in the success of newspaper digitization is the availability of good quality microfilm. Although it is possible to digitize newspapers from an original print copy, this process is very labour-intensive and considerably more expensive than digitalising from film [15].

- Preservation microfilming, carried out in accordance with agreed national standards, produces a set of three films for each title preserved:
- Master negative. This is the film in the camera when the preserved material is photographed. It must not be used except to generate the duplicate or printing negative.
- Printing negative. This is a direct copy of the MNEG. It is used to generate positive or service copies.
- Positive or Service copy. This is the film the library user reads. It is also the format used for copies purchased by individuals or other institutions [16].

Distillation

Distillation processing is performed by Olive Software and includes image analysis, article segmentation, and output to XML. The automatic segmentation process is tasked with recognizing newspaper information objects or entities. Once segmentation has been performed, the print edition is converted to a "Digital Newspaper."

A Digital Newspaper consists of images and XML files. The images are rectangular snapshots which can be used to build up every information object in the newspaper; the XML files record the text, structure and layout of the document. These can be articles, pictures, or ads. It also recognises each entity's internal components (in an article, for example, these include title, subtitle, by line, and body text). All this is done through analysis of page layout geometry and the fonts used on each page [17].

The distillation process was designed to overcome the inherent problems associated with the conversion of scanned images and microfilm as well as the inability of OCR programs to properly read page layout geometry. Distillation is a five-step process: Image analysis, Layout analysis, OCR, Entity building and Output to XML [18].

Archiving Digital Newspapers

Digital newspapers span a diversity of forms. There are newspapers that consist of page images in digital microfilm format, newspapers that have been digitally scanned from analog microfilm and from print (at various image resolutions), encoded text derived from these scanned images (optical character recognition or OCR), and of course born-digital newspapers—often e-prints and web-related text, image, and multimedia files. Since the shift
To 'digital- first' publishing, publishers increasingly have abandoned microfilming of or maintaining their print morgues. Now, newspapers produce a PDF print master to send to their printer [19].

Born-digital content could be provided by the newspaper concerned or harvested by the newspaper's library. The Chronicles in Preservation project (Singapore) studies the preservation challenges of curating these digital newspaper collections. The three distributed digital preservation (DDP) systems are MetaArchive, Chronopolis, and UNTCoda. DDP systems assist with long-term preservation by storing, curating, and validating objects in multiple secure locations. Many different technologies can be used to accomplish this task [20].

The Preservation Technique

The Guidelines include six key functions:

- Inventorying Digital Newspapers for Preservation –How to record what content an organization has and how it is stored
- Format Management for Digital Newspapers –How to identify, validate, and migrate formats
- Metadata Packaging for Digital Newspapers –How to choose metadata formats, export metadata from repositories, and manage the storage of metadata
- Checksum Management for Digital Newspapers –How to generate and monitor fixity information
- Organizing Digital Newspapers for Preservation –How to structure folder hierarchies and names
- Packaging Digital Newspapers for Preservation –How to organize a collection for ingest into a digital preservation system

5. Developments in the Audio-Visual Media

The audio-visual (AV) record of the 20th century is at risk, and digitisation has been a solution, which has created a new problem: preservation of digital AV content. These files have requirements (size or specific formats) that are not adequately addressed by current technology [21]. Achieving this proof of long-term digital preservation depends on two basic premises. First, the information to be preserved is in the digital form. Second, there should be the formulation of proper representation information, which accompanies the raw digital data and is used to interpret it to produce information understandable to its consumers [22]. What do these changes represent for television archives? In a culture that expects broadcast media to be available whenever it chooses, the notion of a video archive takes on new meaning: not as a gatekeeper to accessing older content, but rather as a guardian protecting that content and keeping it vital [23]. The loss of past information and archives for the Mauritian television already spells out very heavily.

The biggest holders of content are broadcasters, and other major holdings are in film museums and other cultural and heritage institutions [24]. Broadcasters vary, but it is common for the computer and technical staff of a broadcaster, and the management who decide and fund technology issues to know absolutely nothing of academic libraries and digital library technology.

Obsolescence of traditional recording formats

With regards to their long-term preservation, audio and video recordings are twins: none of their formats and carriers has been developed for stability over longer periods, and both types of documents are dependent on the availability of equipment which, in the course of technical development, has reached considerable levels of sophistication [25]. Moreover, the ever accelerating pace of technological development has led to ever shorter commercial life cycles.

This is particularly true of magnetic tape based video formats, many of which have been superseded by a follower format after only a few years. Format obsolescence swiftly leads to fading of maintenance support and spare part supply by manufacturers, which leaves even well preserved media irretrievable because of a lack of working playback machines [26].

Traditional Audio Visual Preservation Technique

Traditional preservation of audio-visual tapes or recordings are done in Mauritius by individuals but they might not represent an ideal way of preserving information for the long-term. Quick thinking and improvisation often compensates the scarcity of resources–generic silica gels for magnetic tapes, charcoal as dehumidifier for films, scissors as splicer, isopropyl alcohol inexpensive but lint-free cloth for cleaning audio and videotape heads, preservation winding or (rewinding) magnetic tapes, and air-conditioned rooms as vaults for low-temperature storage are just some of the media centres' inventive measures to keep its preservation efforts at bay [27].

Conversion of Video Tapes to Digital Format

Audio and video preservation is demanding in many respects. Signal extraction from originals needs to meet highest standards, as – differently to digitisation of text and images – originals ultimately will be lost and the digital archival master has to serve as a faithful replica of the original [28]. It is imperative to have modern replay equipment ready for all original formats, to be kept in good working condition by experienced specialists against a fading background of industrial support for spare parts and service. There is general agreement that format obsolescence and the lack of replay equipment is a greater threat than carrier decay. The time window left to transfer contents from analogue and single digital carriers to digital repositories successfully is estimated to be not more than 20 years [29].

The other demanding part of audio and video preservation is the safeguarding of the digital archival master files, which includes data integrity checking, refreshing of data if needed, and migration of data to new preservation platforms in time to avoid loss of information through obsolescence [30]. The challenge of digital preservation as compared to the analogue world is that the maintenance of digital files over generations of storage platforms requires ongoing logistic and financial input in dimensions which exceeds classical carrier-based document preservation by far [31].

Embedded Metadata

Most archives and libraries that manage digital content depend upon the metadata in databases, integrated library systems, and/or digital content management systems. These systems or their extensions also support patron discovery and retrieval of digital content. The charter for the Federal Agencies subgroup devoted to the topic states that embedded metadata plays an important role "in the management, use, and sustainability of digital assets"[32]. "Disaster recovery in the event of the impairment of digital asset management systems depends upon the availability of metadata in standardized formats, including embedded image-level metadata and work-level descriptive, administrative, and structural metadata."

Moving Image Formatting – JPEG 2000

JPEG 2000 is a standard from the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). MXF is a standard of the Society of Motion Picture and Television Engineers, and some refer to it as a container or a wrapper [33]. The growing body of experience with MXF-wrapped JPEG 2000 files means that this is an important target format. There are also other video reformatting efforts, including a trio of activities that entail the capture and storage of uncompressed video streams. One of these is at Stanford University, another at Rutgers, and a third at the BBC [34]. The BBC approach is of special interest because it also employs the MXF container format.

6. A Model for Digital Preservation in Mauritius

From the developments in digital preservation seen in sections 4 and 5, the researcher found it useful to formulate a model regarding how digitisation would be a future requirement for Mauritius in both fields that have been discussed so far.

First of all, there were two distinctive areas that were discussed. Digitisation of newspapers was a separate consideration from that of audio-visual media. The researcher found it unlikely that both types of media (written and audio-visual) could be considered from a holistic perspective although preservation in the form of digitisation impacts both sectors.

The first concern was the digitisation of the written media of Mauritius comprising essentially the local Mauritian press with some five important dailies and some ten weeklies. All such newspapers now have both a hard and a digital format.

Formulation Written Media

Let P_M be the preservation of the written media, the first equation would comprise:

$P_M = A_M$ (archiving traditional paper format) + M_M (microfilming the paper)

$$P_M = A_M + M_M \quad \ldots\ldots (1)$$

Added to this would be:

Digital papers known as e-papers = e

e- Papers would be preserved with a mix of techniques like:

Distillation + Archiving where Distillation would be D and Archiving would be A

Then this would make e (D+A)

The formula for the preservation of the written media could be expressed as:

$$P_M = A_M + M_M + e\,(D+A)\ldots\ldots\ldots\ldots (2)$$

Archiving could be also subdivided as:

i=Inventorying Digital Newspapers for Preservation
f=Format Management for Digital Newspapers
m=Metadata Packaging for Digital
c=Checksum Management for Digital Newspapers
o=Organising Digital Newspapers for Preservation
p=Packaging Digital Newspapers for Preservation
Then $A = (i + f + m + c + o + p)$
This first equation could be summed up as:

$$P_M = A_M + M_M + e\,(D + A\,(i + f + m + c + o + p))\ldots\ldots\ldots\ldots (3)$$

Formulation Audio Visual Media

Let P_{AV} be the preservation of the audio visual media, the first equation regarding such media would comprise:
$P_{AV} = A_{AV}$ (archiving traditional audio visual tape format) + D_{AV} (digitising the audio-visual format)

$$P_{AV} = A_{AV} + D_{AV}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Digitising audio-visual media could be sub-divided into:

Converting analog data to digital data a→d

$$\text{Then } P_{AV} = A_{AV} + D_{AV}(a \rightarrow d)\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

New technologies could be embedding metadata *E* and conversion to the existing international JPEG 2000 format *J*

The formula could be modelled as:

$$P_{AV} = A_{AV} + D_{AV}(a \rightarrow d) + E + J\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

In the long-run when analog systems will be eliminated through conversion to digital format, the equation could be simplified as:

$$P_{AV} = D_{AV}(a \rightarrow d) + E + J\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

and ultimately

$$P_{AV} = D_{AV}(E + J)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

since conversion would be mainly focused on embedding metadata and the use of the JPEG 2000 format.

7. Conclusion

This piece of research has helped to understand how media preservation would be possible in Mauritius in the future. Taking into consideration the lack of attention devoted to preservation in the written media, traditional archiving will prevail and remain an easy and customisable solution. Added to this, the existence of digital format of newspapers can ascertain long-term preservation of information and the possibility for it to be indexed and well stored to meet customer needs. Microfilming is being done by the National Library of Mauritius but will take a long time with an urgent need for resources in every manner; human, financial and physical.

Regarding audio-visual data, it is initially regretted that a lot of analog data in film reels (8, Super 8, 16mm) have been destroyed but video tapes (Betacam SP, VHS, VHS-C) could be converted to standard DVD and onwards to high definition systems. The problem will be to allow such data to be used in a lossless manner. This will require new storage and encoding systems that have been discussed in this research paper.

Preservation is viewed as something that has got something to do with the means by which the documentary heritage is handed down to future generations, while being made available to current users. Past, present and future are the keywords of preservation [35]. Whatever was documented in the past is what present users are benefitting from today, and it is the role of today's users to safeguard the heritage for the future generations to enjoy. Content preservation applies to all societies and, in particular, developing countries like Mauritius where technological improvements could be a cornerstone for the preservation of data in written and audio visual forms.

References

[1] Schultz, M., Skinner, K., 2011. Chronicles in Preservation: Collections Assessment Survey. Educopia Institute.
[2] Library of Congress, *World library body calls for digital preservation of newspapers*, http://www.digitalpreservation.gov/series/challenge/historic_newspapers.html Retrieved 21 December 2013.
[3]Tobey, D, A (2012) *Preserving history Here's how to keep that historic newspaper for years to come*, NYT Regional Newspapers.
[4] Krtalic, M., and Hasenay, D. Preservation of newspapers: Theoretical approaches and practical achievements, University of Osijek, Croatia.
[5] Preservation of Newspapers, http://www.lib.ku.edu/preservation/newspapers.shtml Retrieved 21 December 2013.
[6] Digital Preservation Coalition (2008). "Introduction: Definitions and Concepts". *Digital Preservation Handbook*. York, UK. Retrieved 22 December 2013.
[7]Michael, D (2006). "The long-term preservation of Web content". pp. 177-199. ISBN 3-540-23338-5.
[8]Evans, M., Carter, L., (2008). *The Challenges of Digital Preservation*. Presentation at the Library of Parliament, Ottawa.
[9]Chenganna, A., (2010) Media and the political participation of minorities: The Mauritian Example, University of Mauritius.
[10]Saintville, D. (2009) *La stratégie de sauvegarde et de numérisation des archives de l'Institut National de l'Audiovisuel*, 1999-2015par de mission auprès du Directeur des Archives de l'Ina, Bry sur Marne, France.
[11] (Librarian of Congress, (1997) *Bringing Historic Newspapers to Your Desktop: The National Digital Newspaper Program*, LOC (USA).
[12] (Ide, MacCarn, Shepard, & Weisse, 2002.
[13] Ibid.

[14] Coleman, C. (2012) Preservation Microfilming Guidelines: Newspapers. C.Coleman blog. Retrieved 21 December 2013.
[15]Library of Illinois (2013) Best Practices for newspaper digitisation, USA.
[16] Ibid 14.
[17]Ibid 15.
[18]Ibid.
[19] Library of Congress (2013) National Digital Stewardship Alliance, http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_CaseStudy. Retrieved 22 December 2013.
[20] Skinner, K., Schultz, M., Krabbenhoeft N., and Frederick Zarndt, F. (2013) Chronicles in Preservation: Preserving Digital News and Newspapers, IFLA WLIC Singapore 2013.
[21] Richard Wright, R. (2009) Research & Innovation, BBC richard.wright@bbc.co.uk.
[22] Šimko, V., and Máša, M. (2012) Long-term digital preservation of a new media performance: *"Can we re-perform it in 100 years?"* CIANT International Centre for Art and New Technologies, Prague, Czech Republic.
[23] Rubin, N. (2010) Preserving Digital Public Television:*Is There Life After Broadcasting?* NDIIPP of the Library of Congress, USA.
[24] Ibid 21
[25] Schüller, D (2008) Video Archiving and the Dilemma of Data Compression Special Preservation Projects Phonogrammarchiv, Austrian Academy of Science.
[26] Ibid
[27] Abrigo, C.M, and Abrigo, V.M, (2010) Audio-Visual Preservation: A Perpetual Dilemma Among Selected Philippine Academic Libraries .
[28] Schüller, D (2008) Audiovisual research collections and their preservationSchüller European Commission on Preservation and Access.
[29] Ibid
[30] Ibid
[31] Ibid
[32] Fleischchauer C. (2010) Information Standards Quarterly Spring 2010, Vol 22, Issue 2, ISSN 1041-0031
[33] Ibid
[34] Ibid
[35] Feather, J. (2004). Introduction: principle and policies. In Feather, J. (Ed.), Managing preservation for libraries and archives: current practice and future developments (pp.1-25). England: Ashgate Publishing Limited.