

Planning digitising projects with reference to acquiring appropriate equipment for the project and the quality management process using case studies in South Africa

Antenie Carstens

National Library of South Africa.

E-mail address: antenie.carstens@nlsa.ac.za



Copyright © 2017 by Antenie Carstens. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: <http://creativecommons.org/licenses/by/4.0>

Abstract:

Due to the pressure to digitise collections for public access and preservation, many institutions in South Africa invested large amounts of money to buy large format scanners. Unfortunately very few institutions make use of proper digitising project planning which often results in inaccurate reporting of quantities of the different types of heritage material in their collections earmarked for digitisation. The effect of poor planning, may often lead to wrong decisions when acquiring equipment for the project which can be a waste of resources such as money and labour. The wrong equipment results in under-utilisation as well as inferior digital produced material due to the fact that the equipment was designed for a different purpose. In some cases equipment become redundant due to outdated software or loss of manufacturer support. Often these unfortunate situations can be avoided by proper planning and consultation with experienced institutions or researchers in the field. Lack of communication between organisations in South Africa is often the cause of questionable decisions.

Rapid progress in scanner technology over the past 4 years gave rise to an increase of one-shot systems. Furthermore, the progress towards introducing ISO standards of quality management and the improvement of the delivered quality of the scanners which can achieve these standards, is a major step towards the guarantee of high quality digital content for future access and long term preservation.

The important aspects of quality management (QM) and monitoring are often neglected in projects in South Africa. Locally formulated standards implemented as early as 1999 are not always applicable in new systems anymore. The question is why are project leaders not willing to invest money in a proper QM programme to ensure high quality content for the future?

Keywords: Digitising, Digital Quality Management, Digitising Standards.

Introduction

Most countries have policies regarding their heritage collections. Whether it is the collection, preservation or display, these collections are protected by various policies or national acts. In the case of South Africa, the National Heritage Resources Act, No 25 of 1999, is in place to govern and look after our collections. The Act states that the type of heritage material includes: "books, records, documents, photographic positives and negatives, graphic, film or video material or sound recordings". For the purpose of this paper, we focus on the imaging of visual material for example photographic positives and negatives, books, maps, documents and newspapers, excluding all sound/audio material.

Since the introduction of digital photography, including the electronic means of capturing data by for example the use of scanners, an ever-increasing demand for digitisation of collections was inevitable for the purpose of preservation alternative of originals and for public access of the material. In some instances, digitisation started without proper planning of projects and with a lack of information regarding recommended standards applicable to long term preservation of the files. Apart from scanning standards, other issues such as file formats were not seriously considered or researched. Often the lack of proper planning resulted in the necessary re-scanning of original materials.

There is enough literature and networking available which can be used to facilitate the planning of digitisation projects. From a proper planning approach, it is possible to establish which kind of equipment will be most suitable for the project and particularly suitable for the in-house digitisation process. The choice of equipment can be tricky as new systems become available as technology progresses.

With new developments in technology and a constant increase of delivered quality of the scanning equipment, it is now possible to control and monitor digital quality by means of a digital quality management process. In the case of South Africa, there are only a handful of organisations who apply any quality management processes to ensure consistent quality for the duration of the project. Recent literature reminds us that proper quality management can be beneficial for long-term digital preservation and access (Rieger, 2008:22).

Methodology

The contents of this paper include some information based on the findings of a recent study about digitisation of photographic heritage material, as well as information based on actual observations and personal observations and discussions.

Questions were formulated to establish the level of knowledge that was used to make decisions regarding the digitising projects. Seven heritage institutions in the Western Cape, South Africa were identified and selected and data were collected by means of personal interviews with the curators. Most of the seven institutions had finished their digitising projects by the time of the survey while a few were still occupied with on-going or long term digitisation.

To establish the level of importance of factors regarding file longevity and long term preservation for students in the field of digitising and photography, three educational institutions were identified and questionnaires were used to get their views on specific questions. A similar questionnaire was sent to members of the Photographic Society of South Africa to establish their perception of long-term preservation of visual material. Some of these questions are used here as points of discussion

Networking with similar institutions to the National Library of South Africa at local workshops, conferences and seminars in South Africa, is used on an ongoing basis to establish the current status regarding project planning and the choice and acquisition of equipment for digitising various artefacts.

For the purpose of this paper, the process of digitisation is limited to large format scanners and one-shot systems. Digital cameras are excluded.

Findings

Level of knowledge

An overview of the survey results indicated a level of expertise below average when it comes to identifying material for digitisation, selecting appropriate equipment and the application of standards. Furthermore, decisions relating to image preservation, for example the choice of file formats, were taken without considering the effects of long-term preservation and without consultation or a proper literature study.

In one of the case studies, it was found that the only file format of choice for the photographic negative collection was the jpeg format at an unknown compression. In less than 10 years after completion of the project, water damage as a result of a fire in the archive building, destroy part of the collection. As can be expected the collection needs to be rescanned to be able to preserve the digital files for long-term preservation.

The conclusion was that no formalised project planning was done by any of the surveyed institutions.

Project planning

Literature containing recommended steps for planning digitising projects is widely published or available from several sources. By following the recommended steps, the manager can be sure that all important aspects are addressed. A good example covering all steps to be included in the planning process is proposed by Rieger (2008:16). An extract of the critical steps, which will also give guidance to help with the selection of equipment, include:

- selection of material
- technical requirements
- the digitising process including image processing
- quality control and selection of quality control tools.

Another practical set of guidelines presented by the Public Record Office Victoria of Australia, (Public Record Office, 2011) offer simple steps, guiding managers of digitising projects to follow an appropriate sequence of events. Important aspects are addressed for example:

- skills
- sustainability
- copyright
- privacy
- cultural sensitivity
- creating digital content

In at least three cases in South Africa it was found that expensive digitising equipment was purchased but at the time was not properly utilised either by doing ad-hoc scanning work or just experimenting with the equipment. The main reasons are basically the lack of proper planning of the project in terms of identifying the types of material that needed to be digitised.

In the newspaper digitising project of the NLSA, equipment has been selected that can provide high quality images at a speed that will ensure the completion of the project within the allocated time. However the project does not provide the adding of metadata or public access. Access can only be granted after prior arrangement with the Library. In this regard shortage of funding and specialised staff are the main concerns at present.

Selecting equipment

The recent progress in the design and manufacturing of digitising equipment, gives the project manager a variety of choices. However, care must be taken not to choose equipment that is of such a nature that it cannot be supported in the near future or a type of technology that is already in a phasing-out stage and being replaced by better quality equipment, which may be able to produce high quality at a high output level.

Selecting appropriate equipment for mass digitisation can be a demanding exercise as the variety of heritage materials can be extremely diverse. Therefore, it is seldom possible to digitise all the material with one type of scanner. A proper project plan with an accurate inventory of materials is of utmost importance, which precedes the decision for purchasing equipment. For example, the digitising of the various photographic materials requires very different equipment comparing to the digitising of maps, newspapers and rare books. One of the first choices to consider is the size of the scanning area needed. A proper inventory list will indicate the number of artefacts of different sizes in the collection

and based on the numbers a decision needs to be taken. If for example less than 10% of the collection material is of A0 size, the decision needs to be taken whether an investment of R500 000 (35 0000 euro) more for an A0 type scanner, is justified to accommodate 10% of the material.

To illustrate the difficulty with the process of selecting appropriate equipment, especially if the digitisation project is outsourced to a private company, the following 2015 case study is a good case in point. At the early project planning stage, it became clear that the vendor was not aware of the diversity of the heritage materials with the result that inferior quality and inappropriate equipment was suggested to be used. For some of the artefacts, the design of the equipment was not suitable to produce a high quality image files for preservation, and also risked the possibility of damage to the originals.

Selecting a scanner is only an aspect of the digitising equipment, other important accessories are the book cradles or holders. Care needs to be taken as certain cradles are very restricted in the type of material it can accommodate and offers limited usage. It is not possible to digitise various materials with a single type cradle. For example, V-type cradles may be very difficult to use for scanning of single pages. The same applies to a large size flat table with glass, ideal for flat maps but cannot be used for digitising bound volumes of books.

Output quality increased as technology has progressed in recent years. The elements needed to produce high quality digitised material for example the image sensor, the type of lighting used, lens types and the construction of the scanners itself, all contributed to faster, more effective and consistent high quality output.

Analysis of the personal interviews regarding the selection of equipment, it was found that not enough research was done to establish what would be the best choice for equipment. In one case it was found that a very special negative scanner was purchased from a well-known manufacturer at a price of at least R250 000 (euro 17 350). The scanner could only be operated by an Apple Mac Operating System and ceased working when a Mac upgrade was done and updated drivers were no longer available. Even before the time of a systems upgrade, the manufacturer stopped supporting the scanner which led to the situation of a total loss of the equipment after very little use. The purchase price may not seem a significant amount, but it should be kept in mind that funding for expensive equipment in South Africa in particular is very limited.

The above case is not unique as it was established recently by means of a personal interview with a program manager at a well known heritage institution in South Africa, that three A0 scanners were purchased but no clear identified projects were in place and that the operating staff are only "practising" on the scanners. In this case the total value of the equipment is at least R3.5m (242 6000 euro). The question arises if the funds were well spend.

In the case of the National Library of South Africa, extensive research was done to establish which equipment could give the library the best value for money for the recently initiated newspaper digitisation project. Although the chosen equipment is primarily suited for the project, it can also be productively used for various other future projects at least for the next four years. Specifications based on the physical sizes of the artefacts, the size of the collection, output speed and quality, were important considerations before a final decision was reached for the choice of equipment.

Application of standards

As more research is being done in the field of equipment design and performance of scanners with an increase of delivered quality, the need arises to deliver content of consistent high quality. High quality content is only possible when system performance can be measured and the generated digital data can conform to required standards which can be validated. By setting standards, manufactures were forced over time to redesign systems which were able to meet the required standards developed over the years. Some of these standards became part of the International Standards Organisation. Some of the important standards are the following:

- 1 ISO 14524 (Opto-electronic Conversion Function)
- 2 ISO 21550 (Dynamic Range: Film scanners)
- 3 ISO 16067-1 (Resolution: Print Scanners)

An often-neglected aspect of the digitisation process and important to give attention to is the ISO 126546 standard which refer to the working environment conditions of scanner operators.

Quality Management

There may be different definitions for quality management of digital collections, but what we are referring to here is concerned with the digital imaging of collections, and can be defined as: “ensuring through a pre-determined management and work procedure that pre-set standards are being met”. By implementing a proper quality management procedure, consistent quality and long term preservation can be guaranteed. Another important aspect that will be avoided is the necessity for rescanning material, which does not conform, to the required quality level. Many institutions adopt the principle of scanning right the first time, avoiding rescanning which is time consuming, costly and the possibility of damaging fragile material.

The introduction and application of newly formulated standards for digitisation can be compared to the introduction of the Q-lab concept by Kodak in the early nineties to illustrate the concept. The Q-lab concept was introduced globally and used a standard of quality management for photographic processing chemicals to process E-6 colour slide films. A basic knowledge of the use of chemicals, following a prescribed sequence of analysis, was introduced by means of an analysing kit, resulted in the accurate monitoring of all chemicals in the process. A control strip was used to precisely measure the condition of each chemical in the processing steps which indicated the corrections needed to keep the process within the pre-determined limits.

With the introduction of the Federal Agencies Digitisation Initiative (FADGI) guidelines in the United States of America by 2011 and in 2012 the implementation of the Metamorfoze standards by the National Library of the Netherlands, a similar approach for digital quality management is now available. All aspects of the digitisation process can be objectively measured and adjusted to reach the required levels by using appropriate, device level -and object level targets. Specially developed software is used to do the analysis of the targets and to print reports that serve as confirmation that the quality management system is functioning. Reports are used to adjust scanner settings to ensure all parameters can be measured and are within the tolerance levels. The purpose of the targets is to accurately measure parameters for example colour registration, sample efficiency, true resolution, correct exposure, accurate colour balance, geometric distortion and noise.

Performance of equipment can vary over time and it is recommended that a quality management processes is in place to monitor the different devices. It must also be noted that each device needs to be tested as each individual device may perform differently. The frequency of tests to verify the performance of the digitising systems will vary according to performance consistency of each device. Frequency may vary between infrequent to a weekly or daily routine basis (Williams, 2010:p14).

The importance of digital quality management in the form of application of standards cannot be underestimated as pointed out by Rieger (2008: 35)

"The possibility of problems caused by equipment failure, digitization operator exhaustion, or use of uncalibrated equipment should be anticipated at the point of digitization, and appropriate measures should be in place to prevent them".

When the principle of "scan right the first time" is part of the planning process, it follows that it is time consuming and costly to do any re-scanning of existing material as a result of inferior quality.

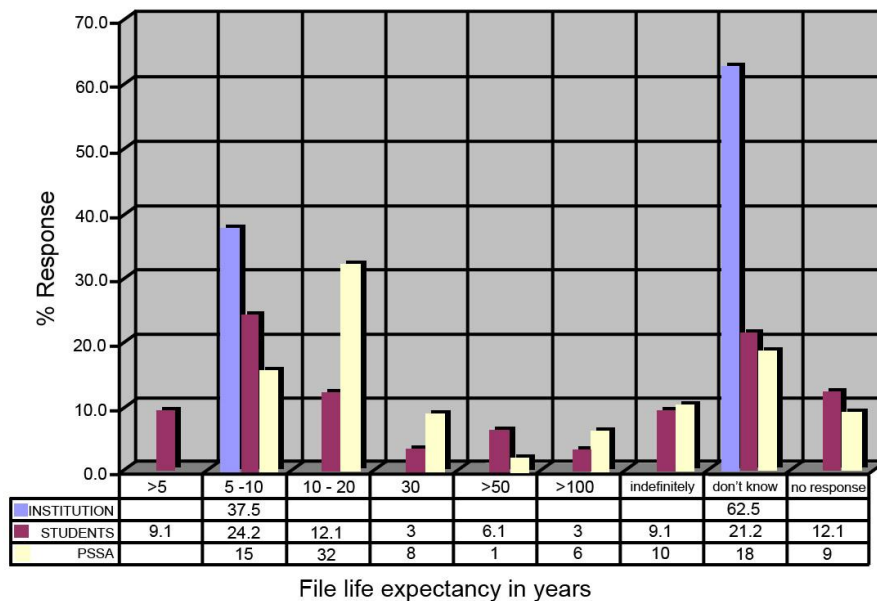
It is often argued that implementing these recommended standards are costly but as Don Williams indicated the cost of a quality management programme versus the cost of high-end equipment might only be 2-3% of the total purchase price of a high-end scanner (Williams, 2014).

Long-term preservation

File format is only one aspect of long-term digital preservation and for the purpose of the survey this was the only question related specifically to preservation. Other infrastructure such as Information Technology (IT) related issues and financial support were not part of the survey. It should however be taken into account that financial restrictions may well be the cause of work deviations that lead to inferior results.

Analysis of the questionnaires with reference to the question of expected file longevity, led to the conclusion that all participating respondents were largely unsure about this aspect of long-term file longevity.

It was expected that the heritage institutions should be considered as professionals, yet life expectancy was considered by 37,5% as between 5 to 10 years and 62,5 "did not know". In the case of the student respondents, results covered a much broader spectrum. Highest percentage 24,4% opted for 5 to 10 years and second high percentage being 21,2% "do not know". There was a 21.1% "no response" (Figure 1).



The response from members of the Photographic Society of South Africa, indicated that 5% of respondents consider the lifetime of a digital file to be between 10 and 20 years. 12% indicated a life of 5 to 10 years, 5% 100 years or more, 8% indefinitely and 14% were not sure.

Conclusion

In a sense South Africa is isolated when it comes to the availability of research and development of new digitising equipment. Another factor is the limited number of vendors who can supply high end digitising equipment. This limitation can easily become a disadvantage to organisations who would like an unbiased opinion when it comes to the selection of appropriate equipment for a given project.

The absence of project plans for digitising projects is evident and the result is ill-informed decisions when equipment is considered. A project plan ensures that a holistic approach to a digitising project is in place whereby all aspects will be addressed and informed decisions can be reached before the

purchase process starts. Networking in South Africa is most important as institutions can prevent spending funds on inappropriate equipment. In fact a standardised project planning framework, which is rigorously applied to every digitisation project which is identified, can go far towards addressing all of the problem areas identified above.

As some institutions are acquiring high-end mass digitisation equipment, quality management must be a very high priority ensuring maximum performance of the equipment. With personal conversations with various institutions, it was concluded that only one heritage institution makes use of a basic quality management program. Most institutions are unaware of quality management as recommended by Metamorfoze or FADGI guidelines. Cost is another factor as local institutions need to cope with severe budget cuts.

References

Public Record Office Victoria, 2011. Just Digitise it. Unnumbered notes, State Government Victoria: pp1-25. [Online]. Available: <https://www.prov.vic.gov.au/sites/default/files/files/Just-Digitise-It%20booklet.pdf> [13 June 2017].

Rieger, O.Y. 2008. *Preservation in the Age of Large-Scale Digitization: A White Paper*. Washington, DC: Council on Library and Information resources; no 141. [Online] Available: <https://www.clir.org/pubs/reports/pub141/pub141.pdf> [Accessed 5 October 2016].

Williams, D. (don-williams@rochester.rr.com). 18 January 2014. Re: version download. E-mail to A. Carstens (antenie@adept.co.za).

Williams, D., Stelmach, M. 2010. Federal Agencies Digitization Initiative - Still Image Working Group. *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files*: pp1-96. [Online] Available <http://www.digitizationguidelines.gov/> [17 September 2016].