

**GUIDELINES FOR USING DIGITISATION AS A SUSTAINABLE  
METHOD OF PRESERVATION AT THE UNIVERSITY OF ZIMBABWE  
LIBRARY**

Mini-dissertation by

**Lovemore Kusekwa  
(11319390)**

Submitted in partial fulfilment of the requirements for the degree of

**MASTER OF INFORMATION TECHNOLOGY (B)**

in the

**FACULTY OF ENGINEERING, THE BUILT ENVIRONMENT AND INFORMATION  
TECHNOLOGY**

**UNIVERSITY OF PRETORIA**

Supervisors: Dr. MJ van Deventer & Dr. H. Pienaar

*November 2012*

## **Dedication**

To Tecla and Lincoln

## **Acknowledgements**

The support I got from family, friends and colleagues is appreciated. Thank you to Dr. Martie van Deventer and Dr. Heila Pienaar for the leadership.

## **Abstract**

This mini-dissertation reports on the guidelines for digitisation that were developed for the University of Zimbabwe Library. A literature study and an empirical investigation were done in order to come up with the guidelines. The literature study was primarily used to inform the researcher on those aspects of digitisation for which guidelines could be developed. A qualitative research design was adopted in this study. Semi-structured interviews were used to collect data from five South African institutions that are involved in digitisation. Digitisation managers from the five institutions were interviewed in this study. The data from the interviews and the literature study were used to come up with the proposed guidelines for digitisation at the University of Zimbabwe Library. The guidelines that were developed as a result of this study specifically address digitisation standards and best practises, technologies for digitisation, costs and sustainability of digitisation, intellectual property issues, steps in digitisation, challenges in digitisation and the selection of materials for digitisation. Although some of the respondents expressed that digitisation can be used as a sustainable method of preservation, this research could not demonstrate that current practices would be reliable in the long term.

## Contents

Dedication .....	i
Acknowledgements .....	ii
Abstract .....	iii
List of Figures .....	vii
List of tables .....	viii
Chapter 1 .....	9
1. Research in context.....	9
1.1. Introduction.....	9
1.2. Research question .....	10
1.2.1 Sub-questions .....	11
1.3. Scope and limitations .....	11
1.4. Justification for the research /Rationale for the study .....	12
1.5. Overview of the literature .....	12
1.6. Research Methodology .....	13
1.7. Value of the study.....	13
1.8. Clarification of key terms .....	14
1.9. Division of chapters .....	14
2. Literature Review .....	16
2.1. Introduction.....	16
2.2. Best practices and standards for digitisation.....	17
2.2.1. Paper based/ printed documents .....	18
2.2.2. Image quality standards .....	18
2.2.3. Films.....	19
2.2.4. Metadata standards.....	20
2.3. Technologies .....	21
2.3.1 Digital cameras.....	21
2.3.3 Scanners .....	22
2.3.4 Film scanners .....	23
2.4. Costs associated with digitisation.....	23
2.5. Steps for digitisation .....	24
2.6. Intellectual Property Issues .....	26
2.7. Selection Criteria .....	28
2.8. Conclusion/ summary of the findings .....	29
3. Research Methodology .....	30

3.1.	Introduction.....	30
3.2.	Goal of the empirical study.....	30
3.3.	Problem definition phase.....	30
3.4.	Research design.....	31
3.5.	Target group and sampling.....	33
3.6.	Data collection method.....	34
3.6.1	Advantages of semi-structured interviews.....	34
3.6.2	Disadvantages of semi-structured interviews.....	35
3.7.	Development of data collection tool.....	36
3.8.	Data analysis and interpretation.....	40
3.9.	Conclusion.....	41
4.	Data Analysis and interpretation.....	42
4.1.	Introduction.....	42
4.2.	Data analysis.....	42
4.2.1	Digitisation standards and best practices.....	43
4.2.2	Selection criteria.....	45
4.2.3	Technologies for digitisation.....	46
4.2.4	Costs of digitisation and Sustainability issues.....	51
4.2.5	Intellectual property rights issues.....	53
4.2.6	Steps in digitisation.....	54
4.2.7	Challenges in digitisation.....	56
4.2.8	Other issues for consideration.....	57
4.3.	Guidelines for digitisation at the University of Zimbabwe Library.....	58
4.3.1	Guidelines on digitisation standards and best practises.....	59
4.3.2	Selection guidelines.....	60
4.3.3	Guidelines on technologies.....	62
4.3.4	Guidelines on costs and sustainability.....	62
4.3.5	Intellectual Property Issues.....	64
4.3.6	Steps in digitisation.....	64
4.3.7	Challenges in digitisation.....	66
4.4.	Conclusion.....	68
5.	Conclusions and Recommendations.....	69
5.1.	Introduction.....	69
5.2.	Conclusions and Recommendations.....	69
5.2.1	Best practices and standards.....	69
5.2.2	Selection of materials.....	70
5.2.3	Technologies for digitisation.....	70

5.2.4	Cost of digitisation .....	71
5.2.5	Intellectual property issues .....	71
5.2.6	Steps in digitisation .....	72
5.2.7	Other issues for consideration.....	72
5.3.	Future and Further studies .....	73
	References.....	74
	Appendix A – Interview schedule.....	80
	Appendix B – Copyright Clearance Template.....	84

## List of Figures

Figure 1: Bookdrive DIY .....	21
Figure 2: Digitising Line (DL 3000) .....	22
Figure 3: HP Scanjet N9120 Document Flatbed Scanner .....	22
Figure 4: Plustek OpticFilm 7400 Scanner .....	23
Figure 5: Scanning resolution in use by the five institutions .....	43
Figure 6: File Formats for preservation.....	44
Figure 7: Feeder scanner used at Institution A .....	47
Figure 8: Flatbed Scanner at Institution A.....	47
Figure 9: Flatbed Scanner at Institution B.....	48
Figure 10: Flatbed Scanner at Institution D .....	48
Figure 12: Suprascan 10 000 scanner at Institution D.....	49
Figure 11: Bookeye 4 Scanner at Institution E.....	49
Figure 13: Book Drive DIY scanner at Institution A.....	50
Figure 14: Nikon Film scanner at Institution D .....	50
Figure 15: Challenges in digitisation .....	57



## **List of tables**

Table 1: Rationale of questions asked.....	37
Table 2: Digitisation standards for the University of Zimbabwe Library .....	59

# Chapter 1

## 1. Research in context

### 1.1. Introduction

Libraries need to preserve important information so that users can continue to have access to the information. Digitisation is seen as “an exciting preservation option while providing unparalleled access available to all” (Berger 1999). This research came up with guidelines for digitisation as a sustainable means of giving access to ‘at risk’ material. This step would assist in preserving selected material from overuse at the University of Zimbabwe Library.

The mission of the University of Zimbabwe Library is to provide access to scholarly information resources required to meet the learning, teaching, research and service needs of the University of Zimbabwe and the vision is to maintain the University of Zimbabwe Library as the leading academic library in the country and in the region. The library seeks to achieve its mission “by building, organising, maintaining, preserving its information resources” and “by providing access to online databases available elsewhere among other things” (University of Zimbabwe Library, 2012). Digitisation as a sustainable method of preservation of information resources can, therefore, be seen as important aspect of the institution achieving its mission and vision as this will ensure enduring access to the information resources. However, digitisation can appear to be complicated and requires guidelines for the processes to follow and the resources to be allocated. This research was done to ensure that guidelines are established that would place the institution in a position where it can successfully embark on its digitisation initiative.

It has been recorded that when it comes to preservation ‘no one strategy is appropriate for all data types, situations, or institutions’ (Gbaje 2011). The guidelines have been established to digitise ‘information materials’ that are currently housed in the library’s special/rare collections. The ‘information materials’ held in the library are available in a range of formats which include print materials, photographs, video cassettes, audio cassettes and microfilm. The items are of historical significance and are written by local authors, illustrators or editors. Access to materials housed in the Special Collections is restricted because of the significance and nature of the items in the collection. This implies that users of the library do

not enjoy as much access as they would like. This is not a unique situation because as Novara (2010) points out: “researcher demand is often a major impetus in the creation of digital surrogates of special collections materials”. Digitisation of materials in the special collections at the University of Zimbabwe Library to preserve and provide electronic access to the materials would ensure that users of the library and the international community will have access to the materials while the original items would incur no further damage through use.

The library also provides access to electronic information sources through subscription and provides access to information sources through its institutional repository. The materials currently held in the institutional repository include journal articles and conference papers by University of Zimbabwe employees.

Many libraries are already involved in digitisation activities (Rieger 2008; JISC 2009; Jones 2004; Lampert & Vaughan 2009 and Liu 2004) among others. The increasing number of libraries involved in digitisation activities means that the guidelines for digitisation as a means of preservation were inspired by the many cases from other institutions that have already begun the processes of digitisation of information materials for preserving and giving access to the materials. The abundance of literature on digitisation also helped the author in selecting criteria to investigate sound guidelines for digitisation at the University of Zimbabwe Library. The literature review was used as background to empirical research that was conducted at five South African institutions that are already digitising their material.

The author’s interest in digitisation was inspired by the presentation by Van der Merwe (2011) on digitisation and migration activities at CSIR during the 4<sup>th</sup> African Conference for Digital Scholarship and Curation; digitisation work at the University of Pretoria Library Services during a tour of the institution in 2011 and digitisation at University of Wisconsin Milwaukee Library (Golda Meir Library) in 2011. These experiences proved to the author that digitisation could be a complicated and challenging process that requires careful planning for it to be successful. It is against this background that the author embarked on research to come up with the guidelines for digitisation at the University of Zimbabwe Library.

## **1.2. Research question**

The research aimed at answering the following question:

How can the University of Zimbabwe Library embark on digitisation as a method of preserving its special collection of information resources?

### **1.2.1 Sub-questions**

The research aimed to answer the following sub-questions:

- What are the best practises and standards that should be used for digitisation?
- Which technologies should be used for digitisation and digital preservation and at what cost?
- What are the steps for digitisation for preservation and for access?
- What are the intellectual property issues to contend with?

### **1.3. Scope and limitations**

The research was focused on establishing guidelines for the digitisation of selected monographs, handwritten manuscripts, theses, dissertations and printed photographs which are based in the 'Special Collections' section of the University of Zimbabwe Library. Most of these materials are rare. This means that the materials could prove difficult or even impossible to replace once they are lost; or when their physical condition deteriorates due to overuse. The limitation of this research is that it is only concerned with the digitisation of selected materials in the special collections of the library. This is a limitation because there may be other information materials besides those in the special collections that deserve to be digitised (for example, maps) and would require separate guidelines for digitisation because of the differences in formats.

As a further limitation it is necessary to acknowledge that the rate at which technology is evolving is overwhelming and brings about a lot of uncertainty (Munir & Phillips 2002). Digitisation and preservation technologies are not immune to changes. It would not be surprising that new solutions to the problem that this research is trying to solve may come up as soon as the research paper is published. As a result, the technology related conclusions and recommendations that have been made may be valid for a limited period of time.

#### **1.4. Justification for the research /Rationale for the study**

Some of the materials that need to be digitised and preserved for long-term use are unique, rare or fragile; hence, they urgently need to be digitised for long-term access and use. It is important to plan for digitisation and preservation especially because there are associated costs that would have to be justified for the investment in digitisation (Eden 2001). This necessitated that a study be done to come up with guidelines for digitising as a means of preserving the information resources while also giving access to the information resources. Without the study, the implementation of the digitisation and preservation activities could prove to be challenging and daunting.

#### **1.5. Overview of the literature**

There is a considerable volume of literature on digitisation and digital preservation that the researcher used in this research. Eakin, Friedlander, & Schonfeld (2006) offer “A selective literature review on digital preservation sustainability” which was quite useful in determining the requirements for sustainable digital preservation initiatives. Bailey’s (2011) “Digital curation and preservation bibliography 2010” also presents an extensive collection of literature grouped according to the following categorisations:

- General Works about Digital Curation and Preservation
- Digital Preservation Copyright Issues
- Digital Preservation of Formats and Materials
- Digital Preservation Metadata
- Digital Preservation Models and Policies
- Digital Preservation National and International Efforts
- Digital Preservation Projects and Institutional Implementations
- Digital Preservation Research
- Digital Preservation Services
- Digital Preservation Strategies
- Digital Repository Digital Preservation Issues

The literature on digitisation projects that libraries around the world have embarked on is also abundant. It is relatively easy to draw lessons from successful digitisation initiatives. The successful initiatives mentioned in the literature provided useful insights for the

researcher. Examples of digitisation projects in the literature are: Jones (2005); van Deventer (2009); Lampert & Vaughan (2009); Iwhiwhu & Eyekpegaha (2009); Kupfer (2010) and Novara (2010) among others.

Literature resources published in 2011 and 2012 which are not mentioned in the sources above, like (Van der Merwe 2011); Bia et al. (2011); Huwe (2011); Balas (2011); Minow & Smith (2011), Vrana (2011) and Brindley (2012), among others, have also been used in the research. The literature study is presented in Chapter 2.

## **1.6. Research Methodology**

A thorough and critical review of literature was conducted in order to enhance the author's understanding of digitisation and long term preservation. The answers to the research question and the sub-questions were sought from the interpretation of the literature on digitisation and the interpretation of data gathered through an empirical study. The knowledge and insight that the author gained through visits to University of Pretoria Library Services and Golda Meier Library in 2011 provided the researcher encouragement that data could be gathered through interviewing experts in libraries that are involved in digitisation.

The empirical study was done by investigating the actual activities at five institutions that are involved in digitising their collections. Physical visits were done to the institutions and semi-structured interviews were used as data collection instruments to collect information from the respondents from the institutions. The respondents that were selected for the study are digitisation managers who have experience and expertise in the entire digitisation process. A detailed methodology for this research that outlines the research design, problem formulation, selection of the target group and the data collection instruments used is presented in Chapter 3.

## **1.7. Value of the study**

The study resulted in guidelines that the University of Zimbabwe Library could use when selecting information resources to digitise in its quest to give access to rare materials. Other university libraries in Zimbabwe can also adapt the guidelines and implement digitisation as a sustainable way of giving access to selected collections at their institutions. However, it must always be remembered that "different ... organizations have different standards and

culture” (Bailey-Hainer & Urban 2004), hence, the need to study these differences and come up with guidelines for use within the University of Zimbabwe context.

### **1.8. Clarification of key terms**

**Digitisation** – is the process of converting analogue materials to machine-readable format for the purpose of providing electronic access and/or preservation (Lampert & Vaughan 2009). Scanning analogue materials such as books and converting them to electronic format is one of the most important tasks in the digitisation process that libraries do. Digitisation is incomplete if the digital objects have not been described fully so that they can be easily identified and retrieved in the future.

**Preservation** – is the process of keeping library materials or information contained in the information sources safe so that they can be used for a long time (Hodge 2004). Digitising information materials is one of the ways of preserving the information contained within the information sources as it involves a change of format of the information source.

**Digital Preservation** – Gbaje (2011) points out that digital preservation involves both the physical and intellectual control of electronic records so that they can be used and accessed over time without compromising the authenticity of the records. Digital preservation is an important aspect of digitisation as information materials are digitised as a means of preserving the content of the information materials. It is important to note that long term accessibility is currently receiving much research attention. Long term preservation constraints will not be discussed here.

### **1.9. Division of chapters**

This mini-dissertation comprises five chapters. The first chapter is an introductory chapter that gives background information about the research. The chapter gives background information about the University of Zimbabwe Library and the materials that could potentially be digitised. It also identifies the research question and the sub-questions; outlining the scope and limitations of the research; justification for the research; overview of the literature; research methodology; value of the study; and clarification of key terms.

Chapter 2 focuses on a literature study that aimed at helping the researcher in answering some of the sub-questions of the research. Literature on guidelines and action plans of other institutions that have digitised or are in the process of digitising was also studied. The literature study helped the researcher to formulate the right kind of questions to ask during the subsequent empirical study.

Chapter 3 reports on the methodology that was adopted to gather information on how some institutions in South Africa are digitising information materials and how they are sustaining their digitisation initiatives. The chapter also outlines the approach that was used in conducting the research. The problem definition phase, research design, data collection methods, the selection of research location, the sampling techniques used and the methods used to analyse and interpret the data are also discussed in Chapter 3.

Chapter 4 focuses on data analysis and interpretation of the data collected from the different cases that were studied. It goes on to present the actual digitisation guidelines for the University of Zimbabwe based on the findings from the empirical study and the literature study that were completed by the researcher. The guidelines given represent the interpretation of the data gathered through the empirical study and the literature study.

Chapter 5, which is the final chapter, summarises and provides conclusions and recommendations on digitisation as a sustainable way of preservation. Recommendations on how other academic libraries can adopt the guidelines to their environments are made in this chapter. The chapter also suggests future studies that could be done that are related to this study.



## **Chapter 2**

### **2. Literature Review**

#### **2.1. Introduction**

Literature on digitisation shows that digitisation can be used as a means of preserving fragile and/ or rare information resources (Berger 1999, Ballard & Donald 2007 and Vrana 2010, among others). Digitising rare or fragile information materials means that libraries can use the digital copies to provide access and provide a means of saving the original materials from overuse. Berger (1999) acknowledges that literature provides an indication that digitisation is 'an exciting preservation option' and also makes it easy to provide access to many users. Ballard & Donald (2007) are also of the opinion that digitisation seems to be 'an effective insurance policy' while Vrana (2010) sees digitisation as 'the only method to safeguard fragile materials' or assist in the preservation of materials. Many libraries around the world have digital collections and they are becoming an important part of library collections (Vrana 2010). A lot of information sources point to the fact that many libraries are already involved in digitisation activities (Rieger 2008; JISC 2009; Jones 2005; Wentzel 2006; Lampert & Vaughan 2009 and Liu 2004, among others). Although Novara (2010) points out that researcher demand is often the impetus for digitising special collections and Kupfer (2010) points out that digitisation is performed with access and preservation in mind, there is evidence that 'one of the main reasons for the initiation of projects of digital collection development is preservation' (Vrana, 2010) and that digitisation has the ability to expand the scope of digitisation activities (Vrana 2011).

Given that digitisation is taking centre stage as a preservation option for many libraries around the world, it becomes an imperative to come up with a set of guidelines on digitisation with the preservation of fragile and/ or rare material (special collections) in mind. Hasenay & Krtalic (2010) argue that the preservation of digital materials is more complicated than the preservation of analogue material and requires 'active management' because of the frequency of changes taking place. However, Hasenay & Krtalic do not indicate which changes are taking place. There is so much literature available on digitisation that van Deventer (2009) advises against 'reinventing the wheel' and to learn from the digitisation

efforts from other institutions that have successfully implemented digitisation projects. It is, therefore, necessary to find answers to the various questions that are raised when a library plans to embark on a digitisation project. Many of these answers can be found in the literature. The next section will provide some insights into the documented current best practices. The intention is to benchmark findings from literature against actual practice. The findings from the benchmark exercise are reported in chapter 4 of this report.

## **2.2. Best practices and standards for digitisation**

According to Liu (2004) the practices, standards and guidelines that are being used by institutions that are digitising or have digitised their collections differ from project to project. It is, therefore, an imperative to have standards for executing digitisation efforts and to also have policies on the standards, preservation and selection of materials to be digitised (Liu 2004). Groenewald & Klapwijk (2010) point out that the use of tools such as XENA that is capable of detecting file formats of digital objects and convert them to open formats that can be of assistance in long-term preservation efforts. They, however, acknowledge that no digital file format will last forever but having standards will increase the chance of preserving digital objects.

Nevertheless, institutions should come up with guidelines and standards that best suit their own conditions. Because of the nature of the special collections being curated by the University Of Zimbabwe this research is interested in the digitisation standards and best practices for the following formats:

- Paper based or printed information/ documentation
- Images
- Films

Both metadata and technology standards were investigated and the research also looked at the selection criteria which institutions that are digitising or have digitised their collections, have employed.

Gbaje (2011) advises that standards should be preferably open and widely available. File formats, for which there is an increased likelihood of stability and longer term support, should be supported. Standards relate to resolution, file formats for the digital images for access by users, bit depth; file formats for optical character recognition (OCR) conversion and also file

formats for the master copies of the digitised images. These standards are crucial as they enable libraries to mount digital collections that are 'compatible, cross-searchable and easily integrated' (Liu 2004). The next section looks at these standards in more detail.

### **2.2.1. Paper based/ printed documents**

Chapman and Kenney (1996) found the 600 dpi, 1-bit, text setting to be ideal for text documents that need to be converted to searchable text as this setting produces an accuracy rate of 99,3%. Berger (1999) points out that scanning at 600 dpi is for archival purposes and also to ensure that the scanned documents would not become obsolete with improvements in technology.

### **2.2.2. Image quality standards**

The digitised images must be authentic, complete and accessible (United Nations ARMS 2006). Gbaje (2011) is of the view that the use of open standards, widely accepted or agreed standards ensures stability and long-term support for digital images.

Novara (2010) points out that as a curator, one should heavily rely on technical experts in order to fully understand the standards that can be used for capturing, providing access and storing images for preservation. The image quality standards vary from project to project and also depend on the type of material that is being digitised (photographic materials, archival documents, printed texts) and the purpose of the digitisation (Riley 2004).

Ballard & Donald (2007) point out that TIFF format is generally used as the archival standard. The following standards are widely used for photographic materials and archival documents (Riley 2004):

- **Photographic materials**

- **Master files:**

- Pixel dimensions: long side of 3000 pixels

- Resolution: sufficient to achieve desired pixel dimensions

- File format: uncompressed TIFF, Intel byte order

- Bit depth: 24-bit colour, 8-bit greyscale

**Access files**

Riley (2004) suggests the following for photographs meant for user access:

Pixel dimensions: long side of 600 pixels

Resolution: 72 ppi

File format: JPEG

Bit depth: 24-bit colour, 8-bit greyscale

- **Archival documents**

**Master Files:**

Pixel dimensions: dependent on size of original

Resolution: 300 ppi

File format: uncompressed TIFF, Intel byte order

Bit depth: 24-bit color, 8-bit greyscale

**Access Files:**

Pixel dimensions: long side of 600 pixels

Resolution: 72 ppi

File format: JPEG or GIF

Bit depth: dependent on file format

Berger (1999) advises that apart from applying these standards, it must be remembered that spots and dirty areas on pages must be removed and images must be sharpened or blurred as necessary so as to achieve the best results.

### **2.2.3. Films**

Puglia, Reed, & Rhodes (2004) point out that “due to photographic limitations of microfilm and the variable quality of older microfilm, it may not be possible to produce what would normally be considered reproduction quality image files.” This view is further supported by the International Federation of Film Archives (2012) who also argue that “there is no theoretical limit to the resolution required to perfectly render an analogue film image, right down to the microstructure of the film grain”. This implies that the standards for scanning films are difficult to set.

#### **2.2.4. Metadata standards**

Gartner (2008) underpins the importance of the standardisation of metadata in digital libraries:

“...the choice of a metadata scheme underpins any such library's ability to deliver objects in a meaningful way, and greatly affects its long-term ability to maintain and preserve its digital assets.”

OCLC/RLG Working Group on Preservation Metadata (2001) and Gartner (2008) point out that the ranges of information that must be included are descriptive metadata, administrative metadata and structural metadata. They both concur that these are essential elements in order to allow for easy retrieval, Curation and rendering of files in a sensible form.

Liu (2004) notes that Common metadata standards used to date are Dublin Core (DC), Resource Description Framework (RDF), Encoded Archival Description (EAD), Text Encoding Initiative (TEI), and Standard Generalised Mark-Up Language (SGML) and its descendents Extensible Mark-Up Language (XML) and Hypertext Mark-Up Language (HTML). Gartner (2008), claims that XML is the ‘standard behind the [metadata] standards’ that are widely used today: Metadata Encoding and Transmission Standard (METS), Metadata Object and Description Schema (MODS) and DC among others. Gartner (2008) also points out that although there are many standards in operation in different libraries around the world, the ‘two that have established themselves most securely in the digital library world are DC and MODS’. DC is particularly popular among libraries because of its simplicity, compatibility with XML and operability with the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH). This observation was originally reported by Guenther & McCallum (2003).

Guenther & McCallum (2003) argue that MODS offers more advantages to libraries because it is a ‘rich metadata schema’ that allows libraries significant scope to capture descriptive metadata, administrative metadata and structural metadata of digital complex objects. They further point out that MODS is interoperable with the MARC 21 standard and compatible with the XML schema for encoding descriptive metadata. Guenther & McCallum (2003) further indicate that METS is another metadata schema that can be used for digital objects. They point out that METS is ‘highly flexible’ and can be used for packaging descriptive metadata and at the same time allow for use and preservation of digital resources. One of the

objectives of the empirical research was to understand which of these standards are being used effectively in the South African libraries that already have successful digitisation programmes.

## **2.3. Technologies**

It is important to look at the technologies that are being used for digitisation. Criteria such as the type of scanner appropriate for your materials, technical capabilities of a scanning system, and budget for scanning equipment and software are key in determining which scanning device a library should choose (Peterson 2005). Flatbed scanners, digital cameras and film scanners are the type of scanners that are commonly used for scanning photographs (Peterson 2005).

### **2.3.1 Digital cameras**

Liu (2004) observes that there is a trend toward using mounted digital camera rather than flatbed scanners as more attention is now being given to the reliability of equipment and software. Peterson (2005) advises that selecting an appropriate scanning device is very important in the digitisation process. The BookDrive DIY book scanner shown below is an example of scanning equipment that makes use of digital camera.



**Figure 1: Bookdrive DIY**  
(Source: BookDrive 2012)

BookDrive DIY “employs powerful digital cameras, mounted either side of the book cradle, to produce sharp, flat-looking images fit for distribution or archiving and ideal for OCR text conversion. Simply place your book facing up on the v-shaped cradle, engage the

transparent platen and press the capture button to shoot both left and right pages.” (Book Drive 2012).

### **2.3.3 Scanners**

Below are some examples of scanners for bound books, text documents and photographs:

#### **Digitizing Line (DL 3000)**

“The Digitizing Line is a scanner for bound works that is equipped with an automatic page turning system. DL 3000 scans up to 2500 pages per hour” (i2S Corporate 2012). See image below:



**Figure 2: Digitising Line (DL 3000)**

(Source: i2S Corporate 2012)



**Figure 3: HP Scanjet N9120 Document Flatbed Scanner**

(Source: Computer Supplies Unlimited 2012)

The HP Scanjet N9120 can be used to scan documents as well as images. It offers “up to 600 dpi optical scan resolution, 48-bit colour, and up to 5000 page daily duty cycle” (Computer Supplies Unlimited 2012).

#### **2.3.4 Film scanners**

Peterson (2005) points out that besides flatbed scanners and digital cameras, film scanners are also well suited for handling special materials like heritage materials in a library. Below is an example of a film scanner that can be used in digitising films.



**Figure 4: Plustek OpticFilm 7400 Scanner**

(Source: B&H Photo video 2012)

#### **2.4. Costs associated with digitisation**

“Technologies for digitisation are expensive” (van Deventer 2009). However, these are not the only costs associated with digitisation. Tanner (2006) observes that sometimes the perceived benefits of digitisation are so huge that people often fail to recognise the costs that will be needed to sustain the digital initiative in order to continue to give access to the digitised materials. Careful planning is, therefore, required in order to ensure that all the costs associated with a digitisation initiative as a whole are known before the process of digitisation can begin (Tanner 2006). Although we can learn from other projects, it is difficult to come up with a formula to calculate the costs of a digitisation effort because each digitisation effort is unique (Hammond & Davies 2009).

Tanner (2006) identifies three factors that he views as key to influencing cost. These factors are; ‘nature of the original item to be digitised’, ‘the digitisation processes and mechanisms



possible' and 'the information, content and delivery objectives to be achieved'. Given that a Minolta PS 3000 scanner that can reach up to 400 dpi resolution costs a minimum of \$5, 000 (Ballard & Donald 2007), the cost of digitisation is certainly high. Besides the high cost of equipment, estimating the cost of digitisation is not an easy task (Bia et al., 2011). Northam (2009) says flatbed scanners range from less than \$100 to thousands of dollars depending on the specifications.

Evens & Hautekeete, (2011) point out that digitisation has proved to be unsustainable because of the costs that are associated with the technologies for digitisation and the digitisation process itself. They, like Bia et al. (2011) also argue that the long-term sustainability of digitisation cannot be guaranteed due to the difficulty in estimating its long-term costs. However, digitisation can still be a useful method of preserving brittle materials or materials that are threatened due to overuse.

Because the literature is inconclusive as to the elements to include when calculating the costs associated with digitisation, the empirical research reported in chapter 4 was used to gain insights into the actual costs to be taken into consideration.

## **2.5. Steps for digitisation**

Chapman and Kenney (1996) are of the opinion that digitisation should begin with project planning and should include selection of materials for digitisation and also the maintenance of the digital files over time. They, however, do not provide greater detail of all the steps that can be followed in digitising collections.

The number of steps that are followed in digitisation of library collections differ from project to project. Liu (2004) indicates that the digitisation process should not require too many steps. Berger (1999) identifies 7 steps while Messmer (2011) identifies 8 steps.

The steps identified by Berger (1999) are as outlined below:

1. Image preparation; preparation of the volumes, issues and pages of the journal for scanning;
2. Scanning pages;
3. Editing images;

4. Using Optical Character Recognition Software (OCR) on images to edit text versions.
5. Creating a searchable database of the text;
6. Linking text to images;
7. Mounting on Web.

The steps are outlined by Messmer (2011) are as shown below:

1. Transport of the original objects to the Scanning Centre
2. Conservational check
3. Preparation for the scanning process
4. Image capture
5. Quality control
6. Indexing and OCR
7. Storage and digital long-term preservation
8. Publication on the web / Access

The steps outlined by Messmer (2011) and those outlined by Berger (1999) are almost similar. The only difference is that Messmer includes transportation of materials to the digitisation site as a step on its own. This step may not be necessary where digitisation is done on-site. Messmer also includes indexing as one of the digitisation steps. This is a very important step as the digitised materials will need to be accessible to the users. Indexing will generate the metadata necessary for the easy location and retrieval of the digital objects.

It appears that the ideal workflow that can be derived from the three sources mentioned above and that address the issue of digitisation workflow can be as follows:

- Selection of materials for digitisation
- Image preparation
- Scanning the pages
- Editing the scanned images
- Optical character recognition (for text files)
- Metadata creation
- Formatting for access and preservation
- Quality control
- Publication on the web

The empirical research was used to confirm what the ideal process is.

## **2.6. Intellectual Property Issues**

The issue of copyright is one that keeps recurring in digitisation literature. Berger (1999) points out that libraries that need to embark on digitisation projects should pay special attention to the issue of copyright so that libraries and publishers could work together to clear any copyright restrictions before materials are digitised. Libraries must go out of their way to ascertain the copyright status of material to be digitised before they begin the digitisation process to avoid violating copyright and intellectual property laws (Jones 2005). Nicholson (2010) considers copyright as a very important aspect of intellectual property because of the legal considerations that have to be made regarding digitisation.

Minow & Smith (2011) emphasise that the task of digitisation is made difficult for librarians because of the restrictions of the copyright laws and advise librarians to know and understand the basic application of the copyright law to avoid trouble. Evens & Hauttekeete (2011) analyse the copyright restrictions of digital materials worldwide and point out that very few countries have clauses in the copyright laws that address digitisation. They give the example of the Digital Millennium Copyright Act in the US as one of the few clauses in the copyright law that allow libraries to digitise materials for preservation and replacement.

In Zimbabwe, the Copyright and Neighbouring Rights Act (Republic of Zimbabwe 2000) points out that:

“The copyright in a work in the permanent collection of a library or archive shall not be infringed by a librarian or archivist who makes a copy of the work—  
... in order to preserve or replace the work by placing the copy in the permanent collection in addition to or in place of it...”.

This provision gives libraries the mandate to digitise rare materials that are endangered. It is, however, not clear in the Act whether libraries are permitted to digitise materials for preservation and make them accessible to library users on the Internet. Libraries might, therefore, need to seek copyright clearance before they can make accessible copyrighted materials that have been digitised. According to Liu (2004) International rules and

regulations are still not standardized. The lack of standardisation means that what might be regarded as 'fair use' in one country may be seen as illegal in another country thus making it difficult for libraries to digitise materials based on the provisions of their country's copyright laws alone.

Minow & Smith (2011) are of the view that the protection that libraries and archives are given by the law gives libraries the opportunity to digitise without fear of prosecution but further point out that this should be backed by clear policy and careful policy. Many libraries that are digitising have, therefore, put measures to deal with copyright, for example the MMM project has employed a system to determine copyright status of materials that is used by the digitising team (Jones 2005).

Fair use is an especially contentious issue in the literature (Lopatin, 2006), but Dames (2005) argues that libraries must exploit the fair use clause in copyright laws to digitise information resources for preservation and also for availing them to their users. Lee (2001), however, points out that libraries now tend to digitise only those materials they have rights over because of the complexity of the copyright laws in various countries. Google's digitisation project's clash with the law may help to prove just how difficult the copyright law can be for libraries wishing to digitise books and make them accessible to users (Green 2010).

Evens & Hauttekeete (2011) and Hughes (2004) recommend that the issue of copyright clearance in digitisation projects in libraries is one that can be dedicated to a specific member of staff. The member of staff will be responsible for contacting the rights holders and negotiating for permissions to digitise materials and make them accessible to users. Lopatin (2006) suggests that any material that is not in the public domain should be negotiated for copyright before it can be digitised. However, Evens & Hauttekeete (2011) point out that 'orphan works' could prove difficult to digitise because of the difficulty in tracing the rights holders.

Although there are issues that may hamper the successful implementation of digitisation projects as seen in the literature reviewed above, libraries can be able to put in place measures that enable them to digitise materials. Having dedicated staff to clear copyright on materials that are still under copyright and focusing on digitising materials that are out of copyright or are in the public domain are also some of the measures that libraries can employ to avoid copyright infringement. It is also likely that libraries can avoid copyright infringement by providing access to digitised materials to only bona fide users of the library.

## **2.7. Selection Criteria**

Librarians should set guidelines for the selection of materials for digitisation and must ensure that the materials that are selected for digitisation are valuable and will not cost the library a lot of money for little benefit (Vogt-O'Connor 2000). Copyright restrictions and the difficulty of copyright clearance usually mean that digitisation is restricted to materials that the library owns rights over or on items out of copyright (Lopatin, 2006). Evens & Hauttekeete (2011) also point out that the length of the rights clearance process is often a deterrent to libraries, hence most libraries focus on digitising materials that are out of copyright.

Hazen, Horell & Merrill-Oldham (1998) suggest that materials can only qualify for digitisation if their intellectual quality is high, are being heavily used, digitisation will result in increased intellectual value of the material and if giving electronic access to the digitised material will lead to enlightenment. Vogt-O'Connor (2000) argues that the selection of materials for digitisation should be an easy process for librarians, archivists and curators who are known to be familiar with the selection process. The empirical study involved asking digitisation managers how the selection of materials for digitisation is done in their institutions. The lessons drawn from the study were used to come up with the guidelines for selection of materials for digitisation at the University of Zimbabwe Library.

Selection of materials for digitisation is done mainly because the digitisation process is an expensive process so it becomes difficult for libraries to digitise everything and also because of the legal issues that may arise as a result of digitising materials that are protected by copyright (Vogt-O'Connor 2000). She also provides a comprehensive checklist that can be used for evaluating materials to be selected for digitisation where she includes such issues as mission statement, collection development statement, privacy, copyrights, authenticity, sensitivity, visual accuracy, stakeholder requests and donor restrictions as among the selection criteria that can be used when digitising collections. Groenewald (2010) pointed out other important issues to be considered in the selection of materials for digitisation. These include the function of the collection, the target audience, the budget available, and sustainability issues among others. The empirical research, reported in chapter 4 will highlight which of these are being actively implemented by the libraries investigated.

## **2.8. Conclusion/ summary of the findings**

The literature study carried out in this chapter shows that the guidelines for digitisation are available but that guidelines need to address individual circumstances – guidelines set out for one institution cannot simply be transferred to another. Institutions with intentions of embarking on digitisation are encouraged to not ‘reinvent the wheel’ but to learn from the experiences of others that have embarked on digitisation before. However, personal circumstances still need to be taken into consideration.

Image standards quality and metadata standards are well documented, so are selection criteria for digitisation, technologies for digitisation and intellectual property issues. Although the technologies for digitisation are expensive, the potential for and advantages of providing access to information to many through digitisation is considerable. Libraries have special provisions for duplicating materials that are irreplaceable and can apply for copyright clearance for materials that are copyright protected. The following chapter focuses on the methodology that was employed in this study to gather information on how other institutions that are successfully digitising their materials are doing the digitisation.

## **Chapter 3**

### **3. Research Methodology**

#### **3.1. Introduction**

The literature study presented in Chapter 2 provides an overview of what some institutions that have embarked on digitisation are doing. This chapter reports on the methodology that was adopted to gather information on how some institutions in South Africa are digitising information materials and how they are sustaining their digitisation initiatives.

The chapter will also outline the approach that was used in conducting the research. The problem definition phase, research design, data collection methods, the selection of research location, the sampling techniques used and the methods used to analyse and interpret the data will also be discussed in this Chapter.

#### **3.2. Goal of the empirical study**

The aim of the study was to establish guidelines for digitisation as a sustainable method of preservation at the University of Zimbabwe Library. The literature study focussed on reviewing the literature available from other institutions that are digitising their collections. The empirical study focussed on finding out how South African institutions are digitising their collections as a method of preservation. The researcher visited five South African institutions to conduct interviews with digitisation managers with the aim of learning from their experiences and expertise. It was envisaged that the findings from the literature would be complemented by the findings from the empirical study. The empirical study confirmed most of the findings from the literature review thereby giving the researcher a strong position to offer recommendations for digitisation at the University of Zimbabwe Library.

#### **3.3. Problem definition phase**

During the problem definition phase, the researcher formulated questions that the research would seek to answer. This research sought to answer the following question:

How can the University of Zimbabwe Library embark on digitisation as a method of preserving a special collection of information resources?

The literature review in Chapter 2 also provided insight on what questions needed to be answered in order to come up with guidelines for digitisation. It follows that under the broad research question given above, the research aimed to answer the following sub-questions:

- What are the best practises and standards that should be used for digitisation?
- Which technologies should be used for digitisation and digital preservation and at what cost?
- What are the steps for digitisation for preservation and digitisation for access?
- What are the intellectual property issues to contend with?

### **3.4. Research design**

Jha (2008) points out that any discussion of research method is “dichotomized and presented in either a quantitative or a qualitative category because the two paradigms have been assumed to be polar opposites and, among some, even separate and distinct scientific absolutes”. He refers to quantitative research as hypothesis testing research and qualitative research as being “multi-method in focus, involving an interpretive, naturalistic approach to its subject matter”. A qualitative research methodology was used for this study where semi-structured interviews were used to collect data. As Creswell (2009) puts it, the ‘idea behind qualitative research is to purposefully select participants or sites that will help the researcher understand the problem and the research question’.

The most important research designs available to the empirical researcher are the experiment, the quasi-experiment, survey studies and the case study (Pienaar 2001). Petty, Thomson & Stew (2012) identify the commonly used methodologies and methods in qualitative research as grounded theory, case study, phenomenology, ethnography and narrative research. Among the qualitative methodologies, the case study was chosen as the research design as it is the most appropriate design to answer the research questions.



The research was based on five case studies. Five South African institutions, that are digitising their collections, were approached to establish their opinions on the issues mentioned above (section 3.3). The research sought to come up with a guideline for digitisation at the University of Zimbabwe Library based on the digitisation taking place at these institutions.

Gorman & Clayton (1997) define a case study as 'an in-depth investigation' which makes it possible 'to derive the knowledge of the wider phenomenon from intensive investigation'. It was possible to derive knowledge from the case of the institutions that were investigated to come up with the guidelines for digitisation as a method of preserving information materials.

The research adopted what Gorman & Clayton (1997) refer to as 'multi-site case studies'. These were undertaken for the sake of diversity in results and to allow generalisation based on the information collected from the five cases. The results were summarised to come up with the best guidelines for digitisation. The guidelines and practises that are common to most the institutions studied were used to come up with the guidelines for digitisation at the University of Zimbabwe Library.

The five institutions that were chosen for the institutions were chosen because of their well known reputation for digitisation in South Africa. The five institutions included two academic libraries and three research libraries. The institutions were seen as the most appropriate to provide respondents who can give meaningful information to come up with the guidelines for digitisation.

Institution A boasts an impressive range of equipment for digitisation and does digitisation for most academic institutions in South Africa. Institutions that outsource their digitisation send their materials for digitisation at this institution. The knowledge and experience that the institution has acquired from years of digitising materials for different institutions was anticipated to be of great help in providing information necessary to complete this research.

Institution B was chosen because they view digitisation from the point of view of a special library. This institution was seen as appropriate for the study because the guidelines for digitisation at the University of Zimbabwe Library are meant for the special collections. The range of materials at Institution B has similarities to materials being held in the special collections at the University of Zimbabwe Library.

Institution C has just completed a guideline for digitisation which is a framework for institutions managing digital collections. The effort that went into producing that guideline was seen as crucial to this research given that this research aimed to produce guidelines for the University of Zimbabwe Library.

Institution D was chosen because it has arguably the largest digitisation initiative in Pretoria. The knowledge and experience that they have in digitisation are invaluable to any institution intending to embark on digitisation.

Institution E was chosen because it is an institution that has just started their digitisation. It was anticipated that the lessons that they gain to get started in their digitisation and the challenges that they faced would be useful to this research.

### **3.5. Target group and sampling**

South African institutions were chosen because there are a considerable number of digitisation initiatives taking place in South Africa. Zimbabwean institutions can learn from what institutions in South Africa are doing. The research targeted digitisation managers in institutions based in Pretoria, South Africa. Convenience sampling was used to choose the institutions that would be used in the research. Convenience sampling involves using 'participants that are easy to get', their 'availability' and their 'willingness to respond' to questions (Gravetter & Forzano 2009). The five institutions that the researcher chose were not only easily accessible to the researcher but also capable of giving substantial information to produce guidelines for digitisation given the experience that these institutions have in digitisation.

The sampling method falls in the category of 'purposive or judgemental sampling' where the selection of the sample is based on the researcher's 'own knowledge of the population, its elements, and the nature of the research aims' (Babbie & Mouton 2001). The researcher was able to judge that digitisation managers were best placed and have the relevant experience to provide information on the guidelines for digitisation that the research sought to achieve.

### **3.6. Data collection method**

The data collection methods available to the qualitative researcher are interviews, focus groups, observation and documentary analysis among others (Petty, Thomson & Stew 2012). Interviews were the preferred data collection method in this research. Robson (2011) identifies three types of interviews which are structured, semi-structured and unstructured interviews. More specifically, semi-structured interviews were used in this research. Semi-structured interviews involve a few pre-determined areas of interest with possible prompts to help guide the conversation (Petty, Thomson & Stew 2012).

Semi-structured interviews are less formal and allow the researcher and the research subject to further discuss the topic (Kumar 1999) thereby providing in-depth information. The researcher used semi-structured interviews to gather in-depth information with digitisation managers from the five selected institutions. An interview schedule (explained in detail in section 3.7 and shown in Appendix A) that would guide the interview process was prepared. The interview schedule guided the researcher during the interview to know which questions to ask and depending upon the answers further probing was done. The advantages and disadvantages of interviews are discussed in section 3.6.1 and 3.6.2 respectively.

#### **3.6.1 Advantages of semi-structured interviews**

The semi-structured interview was selected as a data collection instrument because of the many advantages that it offers in qualitative research designs. The following advantages were enjoyed as a result of the use of semi-structured interviews as a data collection method:

‘Interviewing is a flexible process’ (Gorman & Clayton 1997). Although the researcher had prepared a set of questions to ask during the interviews, it was possible to ask additional questions based on the responses of the interview subjects. This meant that the researcher was able to gather data that would have been difficult to gather through the use of a structured questionnaire which does not allow the research subject to give additional information. The researcher was able to explain and clarify questions and as well ask respondents to qualify answers which seemed unclear.

The other advantage that semi-structured interviews give to the researcher is that these interviews are capable of extracting ‘unexpected pieces of information’ from the research

subjects (Kumar 1999). The unexpected information can help to enrich the research because there is potential to harness information that the researcher would not have anticipated to be useful for the research.

Compared to a mailed questionnaire, the semi-structured interview produces a better response rate (Kumar 1999). The researcher was able to interview all the research subjects that had been identified. This meant that the researcher managed to get complete responses from all the research subjects. This meant that the research plan was not affected by the lack of responses from research subjects.

Semi-structured Interviews give the researcher a chance to explore the topic in depth and in greater detail (Kumar 1999 & Denscombe 2003). This entails that valuable insights can be gained from the depth of the information gathered. Unlike questionnaires, interviews allow the researcher to actively engage the respondents so that detailed information that is useful for the research can be gathered. Apart from getting in depth and detailed information from interviews, Bailey (1982) points out that the research can ensure that all the questions are answered during the interview.

Interviews also have an advantage because the interviewer is able to observe the nonverbal behaviour of respondents (Bailey 1982). The interviewer, therefore, has a chance to assess the validity of the respondent's answers. The non-verbal behaviour of respondents can stimulate the interviewer to further probe and seek clarification for questions that are not answered satisfactorily.

Denscombe (2003) points out that interviews are 'therapeutic'. Research subjects often enjoy the 'rare chance' to talk at length about their work to a researcher who is only interested in gathering information without being critical. This provides a chance for research subjects to be listened to and provides an opportunity to tell others about their work.

### **3.6.2 Disadvantages of semi-structured interviews**

Bailey (1982) points out that the interviewer may fail to understand or misinterpret the respondent's answer. This will make data analysis a difficult task. As Kumar (1999) points out, information obtained through an interview is difficult to analyse. The responses are usually not standard for semi-structured interviews and data can be unique because of context and experience thereby making data analysis a difficult task (Bailey 1982 & Kumar

1999). The researcher tried to reduce this short-coming by investing a lot of time in analysing the responses to get meaningful information out of the data.

Interviews offer respondents no opportunity to consult records so that they can give accurate information to the interviewer (Bailey 1982). However, in this research the researcher mailed respondents to inform them of the topics that would be discussed during the interviews so that they could prepare and look at their records. This was meant to allow respondents the opportunity to give accurate information.

Interviews are known to consume a lot of time and require the researcher to travel for the interviews (Bailey 1982 & Denscombe 2003). The researcher may have a limited amount of resources to travel for interviews and may thus choose to conduct the research in areas that are only easily accessible. This can therefore create a bias which could affect the results of the research. The researcher ensured that the interviews would not consume a lot of time by notifying the respondents about the topics that were going to be covered during the interviews so that the interview process did not take a lot of time.

Denscombe (2003) argues that interviews can be inhibited by the presence of a recorder. Some respondents may not be comfortable with being recorded so they may tend to be selective of what they say. Further, 'what the [respondents] say may not be what they do' (Denscombe 2003). This brings into question the reliability of data collected through the use of interviews as a data collection method. However, the researcher stressed to the respondents that the recordings will be treated confidentially and that their responses would remain anonymous. This made the respondents feel more comfortable.

### **3.7. Development of data collection tool**

The interview schedule was developed from the research questions of the study. The researcher looked at each question that the research intended to answer and formulated the questions that would seek in-depth information to answer a specific question. For example, the first research question which was aimed at understanding the best practises and standards that should be used for digitisation had the following questions in the interview schedule:

- Which standards do you use in digitisation?

- Are you using the same standard(s) for giving access to digitised materials as those you use to archive the digitised materials?
- Please explain why you chose the standard(s) that you have mentioned.
- Which metadata standard(s) do you use for your digitised materials?
- Kindly explain why you chose the metadata standard(s) that you are using.
- Which file formats do you use for archiving and for giving access to digitised materials?

The full interview schedule (Appendix A) which shows the questions that were asked to answer the research questions is attached. Table 1 below shows the categories in which the questions were grouped into and the reasons why each of the questions were necessary in this study.

**Table 1: Rationale of questions asked**

Category	Question	Rationale
Digitisation Standards	Which standards do you use in digitisation?	To know the various standards that are in use
	Are you using the same standard(s) for giving access to digitised materials as those you use to archive the digitised materials?	To ensure that if the standards for access and archiving are different, the respondent mentions all the standards
	Please explain why you chose the standard(s) that you have mentioned.	To determine if the use of the standards is necessary
	Which metadata standard(s) do you use for your digitised materials?	To ensure that the respondents also talk about metadata standards
	Kindly explain why you chose the metadata standard(s) that you are using.	To know the importance of the standards
	Which file formats do you use for archiving and for giving access to	So that all the formats that are used for archiving and

	digitised materials?	giving access to digital content are known
Selection criteria	Which criteria do you use to select materials for digitisation?	To gain insight into how materials for digitisation are selected
	Please explain why each criterion you have mentioned is important	To know the importance of each selection criterion
Technologies for digitisation	Which technologies are you using for digitisation?	To know which technologies are in use
	Please explain in full all the equipment that you are using for digitisation.	To gain in-depth information on all the equipment in use
	Which factors have influenced the choice of technologies that you are using?	To know the factors that influence the choice of technologies from the respondents
Costs of digitisation and sustainability issues	What are the costs involved in digitisation?	To know all the factors that influence costs in digitisation
	How are the expenses/ costs funded?	To know the different ways that institutions can use to fund their digitisation
	Given the complexities and expenses, do you think digitisation can be used as a sustainable method of preservation?	To know whether digitisation can be used as a sustainable method of preservation based on the opinion of the respondents

Intellectual property rights issues	What are the intellectual property rights issues that you have had to contend with in your digitisation programme?	To gain insight into the intellectual property issues that affect digitisation
	How do intellectual property laws in your country help or constrain you when it comes to digitisation?	To know whether the law hinders or helps digitisation initiatives
	How do you ensure that you do not violate copyright when doing digitisation?	To find out the various ways that can be used to avoid copyright violation
Steps in digitisation	Do you make use of a project-based approach to your digitisation? If not, what approach do you use? Please explain the reasons for your approach.	To know whether project management principles are applied in digitisation
	What steps do you include in your digitisation process?	To know the workflow in digitisation
	Why is each of the steps that you have mentioned necessary?	To be able to judge the value of following the steps mentioned
Challenges in digitisation	What major challenges are you facing in digitising your collections?	To know the major challenges that are faced in digitisation in the targeted institutions
	How do you think these challenges can be overcome?	To find out if the digitisation challenges can be solved
Other issues that should be considered	Are there any other issues you would like us to discuss at this time? Is there anything else you think would be helpful for me to know so	To gain insight into other issues relating to digitisation that the researcher had not



	that I can successfully come up with guidelines for digitisation in my institution?	thought of in the study. It was envisaged that this question would give further information that would be useful in coming up with the digitisation guidelines
--	---	--

### 3.8. Data analysis and interpretation

The data that was collected through the interviews was analysed by first categorising the data according to the topics in the interview schedule. This method of analysing data is what Hsieh & Shannon (2005) refer to as 'directed content analysis' where 'prior research about a phenomenon exists' as opposed to 'conventional content analysis' where the researcher allows the formation of categories based on the data collected to generate new insights. This was possible because the categories had been identified during the definition of the research problem and through the literature review that the researcher did before going into the field to collect data. The data analysis involved comparing and contrasting the data in the different categories from the different cases that were studied. After the comparison, the data was summarised to deduce the guidelines that would be adopted for the University of Zimbabwe Library.

Light weight quantitative analysis of qualitative data (Young 1981; Sandelowski 2000; and Hsieh & Shannon 2005) was also done so that the 'data would be susceptible to more meaningful analysis' (Young 1981). From the quantitative analysis, the researcher was able to group the data collected through the interviews according to their topics and represent the data numerically. For example, from the quantitative analysis it was seen that out of the five institutions that were studied, all five institutions use Dublin Core as their metadata standard. The researcher would, therefore, recommend that Dublin Core be adopted as the metadata standard for the University of Zimbabwe Library because of its widespread use in like institutions that are digitising their collections. Chapter 4 of this study reports on the detail of the data collected.

### **3.9. Conclusion**

This chapter looked at the methodology that was employed for this research. The problem definition phase, research design, data collection methods, the selection of research location, the sampling techniques used and the methods used to analyse and interpret the data have been discussed in this Chapter.

The next chapter looks at data analysis and provides an interpretation of the findings. The presentation, analysis and interpretation will lead to the final chapter where recommendations and conclusions are presented.

## **Chapter 4**

### **4. Data Analysis and interpretation**

#### **4.1. Introduction**

This chapter focuses on data analysis and interpretation of the data collected from the different cases. It goes on to present the actual digitisation guidelines for the University of Zimbabwe based on the findings from the empirical study and the literature study (the literature study is presented in Chapter 2) that were completed by the researcher. This means that the data from the literature review and the data from the empirical study complement each other.

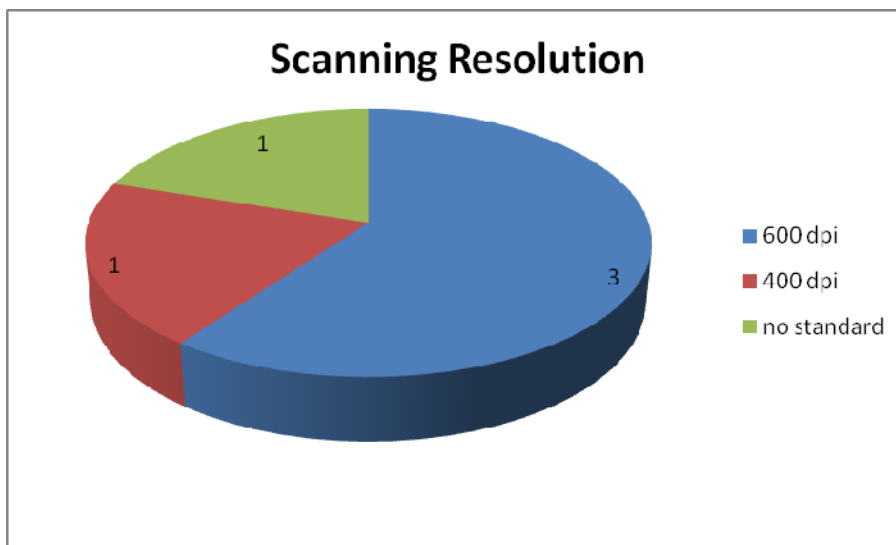
#### **4.2. Data analysis**

The data analysis follows the structure of the interview schedule (see Appendix A) that was used to collect data from the institutions that were selected for the research. The five institutions that were studied will be referred to as Institution A, Institution B, Institution C, Institution D and Institution E. The analysis was done in the order in which the questions were asked during the interviews. Results are reported using the following categorisation:

- Digitisation standards and best practices;
- Selection criteria of materials for digitisation;
- Technologies for digitisation;
- Costs of digitisation and sustainability;
- Intellectual Property Issues;
- Steps in digitisation; and
- Challenges in digitisation.

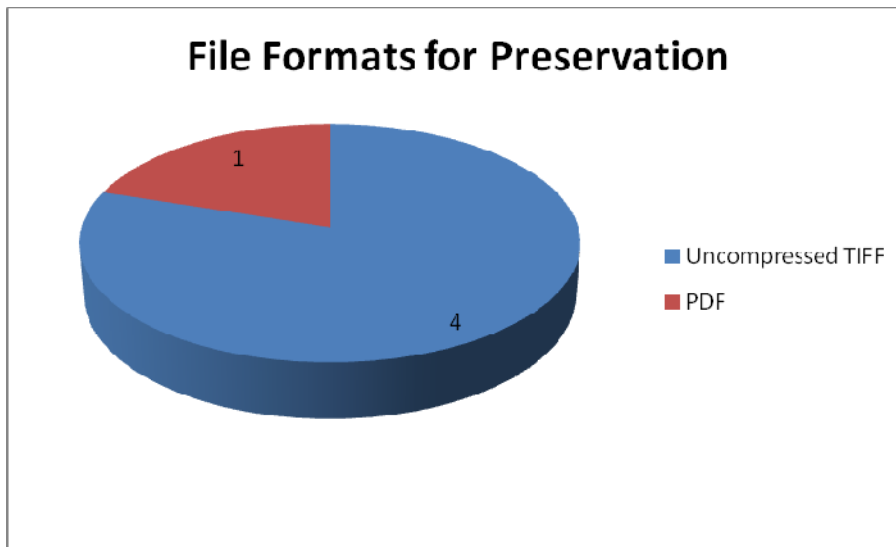
#### 4.2.1 Digitisation standards and best practices

The standards that are used by the five institutions relate to the scanning standards; preservation standards and the metadata standards. The scanning standard that is commonly used is a resolution of 600 dots per inch (dpi). Three out of the five institutions studied use 600 dpi as their standard for scanning with only one institution using 400 dpi. One institution does not use any standard but just looks at the visual clarity of the digitised copy. The visual check is enough for them to decide on the quality of a digitised piece of work. The widespread use of 600 dpi as the scanning standard is in line with the findings from the literature study that was carried out in this research. In the literature, most institutions that are digitising use 600 dpi as their scanning standard – refer to section 2.2.1 in Chapter 2. Figure 5 shows the scanning standards that are used in the institutions that were studied.



**Figure 5: Scanning resolution in use by the five institutions**

From the study it is clear that the standards for archiving digitised materials differ from the standards that are used for providing access to the materials on the web. Tagged Image Format (TIFF) and Portable Document Format (PDF) formats are the formats that are being used among the institutions that were studied. Four of the five institutions studied indicated that they use uncompressed TIFF file formats to archive the digitised materials while only one institution use PDF as their archival standard. Figure 6 below shows that four out of the five institutions that were studied used uncompressed TIFF format as their archival standard; with only one institution using PDF format as their archival format.



**Figure 6: File Formats for preservation**

All five institutions studied indicated that they use PDF format for providing access to the digitised materials on the web. The institutions prefer PDF format for providing access to the digitised materials because it is easy to download. The size of PDF files is considerably smaller than the size of TIFF files. This makes PDF files faster to download from the web. Given that bandwidth in most institutions in Africa is still limited – “a user in Europe enjoys 25 times as much international Internet capacity as a user in Africa” (ITU 2012), PDF files appear to be ideal for providing access to providing access to digitised materials.

The institutions studied also pointed out that PDF format is an open format which can guarantee continuity and provide hope for sustaining the provision of access to the digitised materials. Proprietary formats are not preferred because there is no guarantee that the formats will continue to be in use several years into the future.

The other file format that is used in the digitisation process is Portable Network Graphics (PNG). Institution C and Institution D pointed out that they use PNG format for editing the digitised materials. They pointed out that PNG format is lossless, thus making it the ideal format for editing the digitised materials. Other institutions, however, pointed out that they do not do any editing of the digitised files because they need to preserve the integrity and authenticity of the digitised materials. They argue that the digitised material should resemble as closely as possible the original material that is being digitised.

The five institutions that were studied all use Dublin Core as their metadata standard. Dublin Core is the preferred standard for metadata because of the ease of use and interoperability

with other standards such as MARC 21. Institution B and Institution D indicated that they chose Dublin Core because it is an international standard that is widely used by other institutions around the world. The literature study (Chapter 2, section 2.2.4) also proved that Dublin Core is a very popular metadata standard used in most academic institutions.

#### **4.2.2 Selection criteria**

Institution A is not involved in selection since it only digitises materials for other institutions, while institutions B, C, and E assess the value and evaluate the reason for digitisation to justify selection. Only institution D was able to give a detailed selection process to identify materials for digitisation. Institution D follows the following steps when selecting materials for digitisation:

- **Physical inspection** – the materials are physically inspected to check whether it is possible to successfully digitise the materials. Some materials may be in a very poor state which makes it impossible to digitise them. This aspect was also mentioned by the respondent at Institution A. She pointed out that if the font in a book is so feint that it cannot be read using the human eye, there is very little that technology could do to ensure that the font is visible on the digitised version. Increasing the resolution may help but the problem is that the file size becomes too large to be manageable. Physical inspection is, therefore, meant to identify those materials that could result in a waste of effort so that they could be separated from the materials that are easy to digitise because they are in a good physical condition.
- **Meet with stakeholders** – Meeting with stakeholders is an important step that institutions do in selecting materials for digitisation. This is done to ensure that only materials that the stakeholders deem worthy should be digitised. It is also done to get commitment from everyone involved with the materials being digitised.
- **Check copyright status** – the copyright status of materials is also considered as an important criterion in selecting materials for digitisation. This aspect was also mentioned by the respondents at Institutions B, C and E. All indicated that they give priority to digitising materials that are the intellectual property of the parent institution. All materials that are protected under copyright have to be cleared first before they can be digitised. Four out of the five institutions indicated that they seek copyright

clearance for all the materials that are under copyright. Materials which copyright clearance cannot be granted are not usually digitised. However, Institution D indicated that some materials whose owners cannot be located are digitised. However, they give a clear indication that the copyright owners could not be traced. This gives the copyright owners the chance to claim their work.

- **State the purpose of digitising** – The institutions studied indicated that the purpose for digitising the materials also determines which materials will be digitised. For example, Institution D indicated that if the purpose for digitising is to preserve the original materials by using the digitised copies for providing access, then those materials that are difficult to replace if they are damaged are prioritised. Similarly, Institution B indicated that digitising is the best way of providing access to materials that are in high demand. Digitising materials in high demand reduces wear and tear on the original materials so it would be a legitimate reason to prioritise some materials when it comes to digitisation.
- **Establish the value of the material** – Some materials are more valuable than others. The institutions that were studied indicated that the value of the material is an important consideration when selecting materials for digitisation. Institution B, C, D and E all indicated that the value of the material determines whether a piece of information material is worth digitising. There is no point in wasting resources digitising materials that are not going to be used.
- **Consider the time factor** – Some materials are bulky and will take more time than others to scan. Institution D indicated that those materials that are quick to scan may be preferred to materials that take a long time to scan. This could be because the materials that take a long time to scan will increase the cost of scanning and take a considerable amount of staff time. Materials that can be scanned quickly and made accessible are, therefore, preferred.

#### **4.2.3 Technologies for digitisation**

The research carried out in the five institutions indicates that one type of scanner is not sufficient for digitising library materials. Three institutions out of the five studied have at least a flatbed scanner and a book scanner. Two of the institutions only have flatbed scanners for

scanning unbound materials but they outsource some of their digitisation work that cannot be scanned using flatbed scanners (for instance books and oversized materials). The researcher was given the opportunity to digitally record, (take photographs of) the scanners that are used in the institutions that were studied. The images below show some of the scanners that are used in the institutions that were studied.



**Figure 7: Feeder scanner used at Institution A**

The feeder scanner shown in Figure 7 is used at Institution A to scan materials that are unbound. The scanner can scan up to A4 size paper. This type of scanner is relatively faster than flatbed scanners because it takes in loose papers automatically from the feeding tray.



**Figure 8: Flatbed Scanner at Institution A**





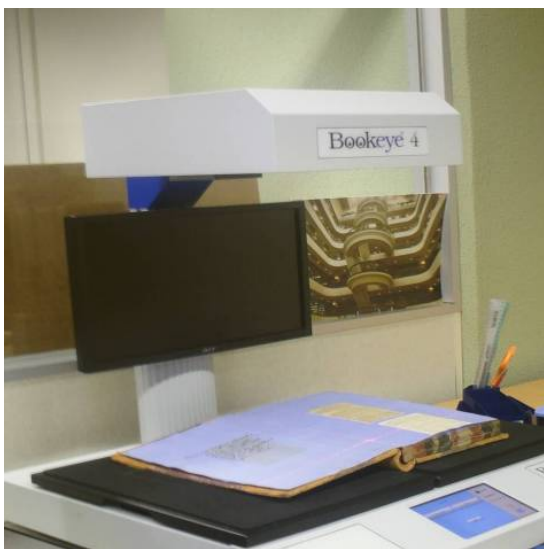
**Figure 9: Flatbed Scanner at Institution B**



**Figure 10: Flatbed Scanner at Institution D**

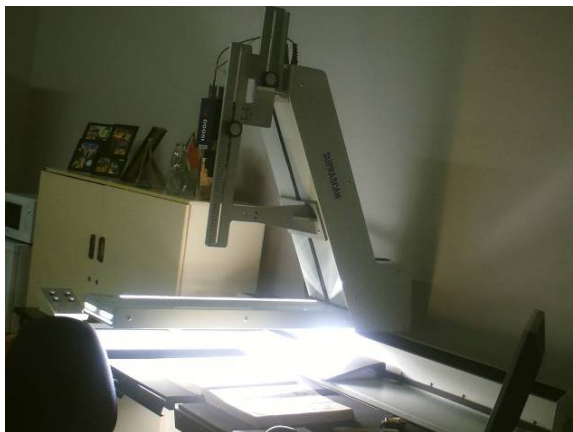
The flatbed scanners in Figure 8; Figure 9 and Figure 10 are used to scan unbound materials. The scanners are capable of scanning basically the same type of materials. The difference is that they are manufactured by different manufacturers. Institutions have different preferences hence the differences in the brands of the flatbed scanners in the institutions studied. The prices of the scanners also differ so institutions choose the brand that suits their budget.

Among the institutions studied the technology that is used for digitising books use overhead scanners. The book scanner at Institution E (Bookeye 4) has a mirror mechanism that is used to capture a book placed on a 120° cradle. Unbound papers can also be placed on the flat surface for scanning.



**Figure 11: Bookeye 4 Scanner at Institution E**

Institution D has a Suprascan 10 000 scanner that uses camera technology. The scanner has a mounted overhead camera that is capable of scanning bound books of different sizes up to size A1. Figure 12 below shows the Suprascan 10 000 scanner.



**Figure 12: Suprascan 10 000 scanner at Institution D**

Institution A uses the Book Drive DIY scanner for scanning bound books. The scanner has two mounted overhead cameras that are used to capture images of books placed on a V-shaped cradle below the cameras. Figure 13 below shows the image of the Book Drive DIY scanner.



**Figure 13: Book Drive DIY scanner at Institution A**

Apart from flatbed scanners and book scanners, the study also found out that film scanners are also needed to digitise microfilm and slides. Among the five institutions that were studied Institution D had film scanners used for this purpose. The image below (Figure 14) shows the film scanner found at Institution D.



**Figure 14: Nikon Film scanner at Institution D**

The choice of which equipment and software to use in all the five institutions has mostly been based on the cost of the equipment to the institution. Each institution indicated that they chose the equipment and software that they could afford but as well the equipment that closely matched their needs as much as possible. The institutions studied indicated that most of the scanning equipment comes with the scanning software installed.

Institution A indicated that the type of materials that need to be digitised also influences the choice of technology to use. For example, an institution cannot choose to digitise fragile materials using a feeder scanner because it can destroy the documents. They also indicated

that different scanners are for different type of documents: Flatbed scanners for unbound materials, overhead scanners for bound books and film scanners for scanning films and slides. This implies that the type of materials that an institution intends to digitise has a direct influence on which equipment to buy in order to successfully digitise materials.

#### **4.2.4 Costs of digitisation and Sustainability issues**

Taking into consideration what was reported in section 4.2.3 above, it is clear that the costs of digitisation are driven by the need to buy equipment for digitisation as well as labour costs. The institutions studied indicated that digitisation is an expensive exercise. These are the two major factors that drive the costs of digitisation. All the five institutions mentioned these factors as influencing the costs of digitisation. Institution A indicated that a book scanner may cost anything between R300 000 (USD34, 050) and R1 million (USD113, 500). Institution B indicated that Adobe Acrobat A software, which is one of the applications used in the digitisation process, may cost as much as R10 000 (USD1, 135) per licence. Any institution planning to digitise materials should, therefore, prepare a budget to cater for the cost of buying equipment and for paying the employees involved in the digitisation process.

Institution B indicated that even if an institution chooses to outsource most of its digitisation activities there are costs that are inevitable. The cost of transporting the materials to the digitisation site and back to the parent institution should not be underestimated. The institution from which the digitisation is outsourced also needs to be paid for the digitisation service. This makes digitisation expensive regardless of whether an institution is doing the digitisation on its own or outsourcing the digitisation service.

Three of the institutions (A, B and D) that were studied said that they manage to sustain the digitisation process because they charge clients the costs of digitising materials. Institution B also indicated that besides passing on the costs of digitisation to the clients, they manage to fund the digitisation by using their annual operational funds from their parent institution because the digitisation has been incorporated in the usual work flow.

Institution C has managed to fund their digitisation initiative through a grant that they were awarded by the Carnegie Corporation while Institution D had managed to get the funding for its digitisation from the project owners (The Hans Merensky Foundation). Institutions intending to embark on digitisation can, therefore, apply for grants to support their digitisation activities.

All five respondents are of the opinion that digitisation can be used as a sustainable method of preserving information but are sceptical that these formats will remain accessible in the longer term. As a result, they offered advice on what they think needs to be done in order for digitisation to be a sustainable way of preserving information. Below are their responses regarding the use of digitisation as a sustainable method of preservation:

The respondent from Institution A said that digitisation can be sustainable but only if people are dedicated to it and do proper planning as well as employing permanent staff to manage the digitisation.

The respondent from Institution B indicated that digitisation is a good way of sharing information while safeguarding the original. She, however, indicated that the long-term sustainability of digitisation has not been proven yet because of technological obsolescence. She also indicated that trying to migrate digital objects for preservation is problematic since some data may be lost during the migration.

The respondent from Institution C was of the opinion that digitisation could be sustainable if an institution adopts a digitisation policy. She also indicated that top management should be committed to investing in digitisation in order for the digitisation to be successful.

The respondent from Institution D indicated that digitised materials can be stored on an archival server. She, however, indicated that hard drives and DVDs have short life spans.

The respondent from Institution E was quick to point out that digitisation is not synonymous with preservation and that long-term planning at institutional level should be done to ensure that digitisation can be used as a sustainable method of preservation. The respondent also pointed out that it is not easy to preserve digital objects.

The above responses show that institutions are still sceptical about using digitisation as a sustainable method of preserving information. However, some think that with proper, strategic long-term planning, dedication, commitment, adoption of policy and adequate financial support, digitisation can offer hope for sustainable preservation of information.

#### **4.2.5 Intellectual property rights issues**

All of the five institutions studied pointed out that they have had to deal with copyright issues in their digitisation initiatives. They pointed out that copyright clearance has to be sought for all materials which are still under copyright and that are not the intellectual property of the digitising institution. The literature study (refer to section 2.6 in Chapter 2) done in this research confirmed that copyright issues cannot be ignored during digitisation projects.

Three of the institutions that were studied pointed out that intellectual property laws can be a hindrance to digitisation. They claim that some materials cannot be digitised because they are protected by copyright. One of the respondents said that she is not sure whether the intellectual property laws are a hindrance or a help in digitisation because they do not have much experience in that area. The respondent from Institution A pointed out that they can digitise anything without any fear of breaking the law. She said that the law only becomes a hindrance when one wants to openly share the digitised material.

There are various ways that the institutions use to avoid violation of the copyright law. All the institutions studied pointed out that copyright clearance is sought in order to digitise and make accessible all materials that are under copyright. The institutions make an effort to find the publisher or author of the works identified for digitisation and ask for permission in writing to digitise the materials. Some materials are, however, difficult to get clearance for because some authors' or publishers' contact details cannot be found. Although clearing copyright is not easy, institutions need to ensure that they do not violate the law.

Institution E pointed out that materials like political posters are just digitised and made accessible even if the copyright owners of the materials cannot be found. They will, however, insert a note on the poster to notify the public that the owners of copyright could not be located. The owners of copyright would, therefore, be required to identify themselves and claim their rights.

Institution A and Institution D also said they avoid violating copyright by digitising materials that are the intellectual property of the digitising institution. They also digitise materials that are in the public domain.

#### **4.2.6 Steps in digitisation**

The sequence and number of steps that are followed, during the digitisation process, vary from institution to institution. The steps that each institution follows are laid out below:

Institution A follows the following seven steps:

- Setting up the scanner to ensure the quality is right
- Calibration of the scanner
- Scan the object as a TIFF file
- Quality control – if the quality is not satisfactory, the material is re-scanned
- Convert TIFF files to PDF files
- Assign metadata
- Upload the object and its metadata to the client server

Institution B embarks on planning for the digitisation before taking any steps to digitise materials. If the document to be scanned is a single document “we simply go to the machine and scan it”. For batches of documents they set the timelines and the digitisation solution for the batch of documents. A decision is made on whether to outsource the digitisation or do it in-house. If a request for digitisation is received, the following steps are taken:

- Go to our archive and retrieve the original publication
- Check whether it’s ok to distribute it and seek clearance if necessary
- Go to the machine to scan
- (If pages are loose, take off staples) and feed it into the scanner - (if bound, flip the pages manually). “We do not do any editing because we do not want to damage the integrity of the document (especially images)”
- Quality control (redo the document until the results are satisfactory)
- OCR upon request (Using Abbyy Software)

Institution C follows seven steps in their digitisation process. The steps are as follows:

- Selection of content and collection
- Checking copyright status
- Set up scanner
- Capture image
- Optimise for web publishing,

- Assign metadata
- Conduct quality control

Institution D provided most detail for the steps that they follow in the digitisation process. The following steps are followed:

- Identification of project, applying selection criteria and clearing copyright
- Basic preservation of original material
- Administration, scanning and archival storing – determine file naming, determine metadata, create databases, quality control, scanning according to international standard, scanning on archival sever and back-up
- Derivating – this process involves copying project from archival server, converting PNG files to a single PDF file, and recognising text using OCR in Adobe Acrobat. File sizes should not be more than 2MB for easy downloading
- Physical and digital preservation – Preserve originals and upload digital objects to UPSPACE

Institution E goes through the following eight steps:

- Assessment of value of material
- Copyright checks
- Document preparation
- Set the scanning resolution
- Scanning
- Quality control
- Metadata creation
- Archiving

A closer look at all the steps that each of the institutions is following in their digitisation initiatives showed that all the institutions basically follow the same steps. Some institutions add a few steps more than others but in principle they all use the same procedures. Some of the institutions were able to outline the steps in greater detail than others during the interviews. However, combining the steps from all the five institutions yields a comprehensive digitisation workflow that can be adopted for use by other institutions. The workflow presented below represents the steps that can be followed when the steps given by the five institutions are combined:



- Selection (assessment of value of material, determine the cost in relation to value)
- Copyright checks (determine status, seek clearance)
- Document preparation (removing staples, replace missing pages, remove binding)
- Scanning (Set the scan resolution, image capture)
- Metadata creation
- Quality control
- Upload to archival server
- Optimisation for web publishing (OCR, Convert TIFF to PDF or JPEG, compress)
- Publish on web

#### **4.2.7 Challenges in digitisation**

There are many challenges that are common to all the institutions that were studied. As shown in Figure 15 below human resources challenges are common to four of the five institutions that were studied. This makes it the most common challenge to digitisation in the institutions studied. Keeping up with technology and the costs of equipment are the other challenges that are prevalent in the studied institutions. Each of the two challenges was reported by two of the five institutions as affecting the digitisation initiatives at the institutions. The poor quality of original materials is a challenge that was reported by one of the five institutions. Poor originals make the process of digitising difficult. For example, a book with torn or missing pages could prove difficult and time-consuming to digitise. Another challenge that was reported by only one institution is getting approval to sustain the digitisation programme. It was pointed out that the long-term sustainability of digitisation is in doubt if management does not commit to sustaining the programme.



Figure 15: Challenges in digitisation

#### 4.2.8 Other issues for consideration

Other issues that could be considered, that had not been discussed during the interviews, are linked to training, the need to have skills for digitisation and the need for institutional commitment to the digitisation exercise. All these issues merely re-emphasise the challenges that need to be dealt with if an institution embarks on digitisation. There is obviously the need to train the human resources that will be involved in the digitisation to have the skills necessary to complete digitisation successfully. The issue of institutional commitment is also another challenge that has already been highlighted as an issue of sustainable digitisation. The digitisation initiatives can only be sustainable if there is commitment from top management to support the digitisation in the long-term.

The other issue that came out of the empirical study, particularly from Institution A is the issue of a Digitisation Management System (DMS). The respondent pointed out that although they do not have one, it is important to have a DMS for managing the digitisation workflow. The DMS would ensure that all the steps in the digitisation process are followed and it can be able to track what has been digitised. This means that duplication of effort can be easily be avoided since the DMS would have all the information regarding material that has been digitised.

### **4.3. Guidelines for digitisation at the University of Zimbabwe Library**

The University of Zimbabwe Library is the oldest academic library in the country and it is arguably the biggest academic library in the country. The library has a large collection of materials. The total holdings of the library include over one million volumes of printed monographs and periodicals. The library serves all the 10 faculties at the university, with each faculty being supported by a faculty librarian.

The information materials held in the library are available in a range of formats which include print materials, photographs, video cassettes, audio cassettes and microfilm. The library provides access to electronic information sources through subscription and also provides access to information sources through its institutional repository. The materials currently held in the repository include journal articles and conference papers by University of Zimbabwe employees. Rare collections, photographs, materials of historical significance and materials written by local authors, illustrators or editors are kept in the special hard copy collections. Access to materials housed in the Special Collections is restricted because of the significance and nature of the materials in the collection. This implies that users of the library do not enjoy as much access as they would like. Digitisation of materials in the special collections at the University of Zimbabwe Library to provide electronic access to the materials would ensure that users of the library and the international community will have unrestricted access to the materials.

The mission of the University of Zimbabwe Library is to provide access to scholarly information resources required to meet the learning, teaching, research and service needs of the University of Zimbabwe and the vision is to maintain the UZ Library as the leading academic library in the country and in the region. The library seeks to achieve its mission “by building, organising, maintaining, preserving its information resources” and “by providing access to online databases available elsewhere among other things” (University of Zimbabwe 2012). Digitisation as a way of preservation of information resources can, therefore, be seen as important aspect of the institution achieving its mission and vision as this will ensure enduring access to the information resources.

The digitisation of the information materials in the special collections requires following a set of guidelines. The guidelines presented below for digitisation at the University of Zimbabwe Library are based on the findings from this research. It is envisaged that using these guidelines is likely to result in a successful and sustainable digitisation programme at the institution. The guidelines are presented under the following topics:

- Guidelines on digitisation standards and best practises
- Selection guidelines
- Digitisation technology guidelines
- Cost and sustainability guidelines
- Intellectual Property Issues
- Steps in digitisation
- Challenges in digitisation

#### 4.3.1 Guidelines on digitisation standards and best practises

The guidelines presented in Table 2 below are proposed for the University of Zimbabwe Library in light of the standards and best practises revealed in the literature study and the empirical study.

**Table 2: Digitisation standards for the University of Zimbabwe Library**

Type of material	Archival standard	Web Access Standard	Metadata standard
Paper based documents	Resolution: 600 dpi File Format: Uncompressed TIFF	Resolution: 200 dpi File Format: PDF	Dublin Core
Images	Resolution: 600 dpi File Format: Uncompressed TIFF	Resolution: 200 dpi File Format: JPEG	Dublin Core

As pointed out in the literature study (section 2.2.3 in Chapter 2), the theoretical limit to the resolution for film is difficult to set. It is therefore recommended that when scanning films, emphasis be placed on producing images that are reasonably clear. However, care must be taken not to produce digital objects that have too big file sizes which might present problems for storage.

The empirical study proved the assertion by Ballard & Donald (2007) that TIFF format is generally used as the archival standard. It is therefore recommended that the University of Zimbabwe Library uses TIFF format for archiving digitised materials. Access to materials is

generally provided in PDF (for text documents) and JPEG (for images). These are the standards that are recommended for the University of Zimbabwe Library.

Dublin Core is the metadata standard that most institutions like the University of Zimbabwe are using. All the five institutions that were studied are using Dublin Core as the metadata standard and the literature survey (Chapter 2) demonstrated that Dublin Core is the most popular and widely used standard. Dublin Core has also been chosen because of its interoperability with the MARC 21 standard that is widely used in describing library materials in most academic libraries.

It is in the interest of the University of Zimbabwe Library to ensure that they also observe the best practises that other institutions that are digitising their collections are doing. The major aim of digitisation is to provide easy access to information materials. Using OCR to ensure that text documents that have been digitised are searchable is mandatory. The full text of digitised documents should be searchable so that the digitised materials can be discoverable even with search engines.

Another best practice that the University of Zimbabwe Library should be doing is the clearance of copyright for all materials that are under copyright. Although the issue of copyright is discussed under guidelines on intellectual property issues, it is imperative to stress that permission to digitise materials that are protected by copyright should be sought before the materials are digitised. In circumstances where the copyright owners cannot be located, the institution should attach a notice to the copyright owners that the copyright owners to notify them that they could not be located before the digitisation was done. The permission to continue to give access to such materials should be sought once the copyright owners have identified themselves.

#### **4.3.2 Selection guidelines**

The criteria that the institutions use in selecting materials for digitisation could prove useful for the University of Zimbabwe Library. Most of the selection criteria that were identified during the literature review (section 2.7, Chapter 2) and the empirical study are recommended for use at the University of Zimbabwe Library.

This study found that the value of materials is an important consideration in selecting materials for digitisation. In the literature study, it was highlighted that Hazen, Horell & Merrill-Oldham (1998) suggest that materials can only qualify for digitisation if their intellectual quality is high and are being heavily used. It is therefore important that the materials selected for digitisation at the University of Zimbabwe represent value for money. Materials that are on high demand and add value to teaching, learning and research should be given high priority in digitisation. As was previously reported; there is no point in wasting time digitising materials that are not going to be used. Meeting with stakeholders may help to determine the value of materials to the users. It is also important that the purpose for digitising the materials be agreed upon by the stakeholders so that everything is clear from the beginning.

As has been highlighted from the data analysis above, physical inspection of materials is an important aspect in selecting materials for digitisation. This is consistent with Vogt-O'Connor (2000); as pointed out in the literature review (section 2.7, Chapter 2) in this study; that materials selected for digitisation should not cost the library a lot of money for little benefit. The University of Zimbabwe should, therefore, take this into consideration so that materials that are in very poor physical condition are not selected for digitisation as this may cost the institution a lot of money to digitise the materials. Materials that are still in good physical condition should be prioritised. Apart from being easier to digitise, materials that are in good physical condition could be prevented from wear and tear by digitising them.

This study has shown that copyright has a bearing on which materials can be selected for digitisation or not. Some of the institutions studied pointed out that they concentrate on digitising materials that are their own intellectual property. As pointed out in the literature review by Lopatin (2006) and Evens & Hauttekeete (2011), copyright restrictions and the difficulty of copyright clearance usually mean that digitisation is restricted to materials that the library owns rights over or on items out of copyright. The University of Zimbabwe should, therefore, choose to digitise first the materials for which copyright belongs to the University of Zimbabwe or are out of copyright. The digitisation should, however, not be limited to the institution's own materials and out of copyright works because copyright clearance can be sought to enable the institution to digitise materials that are under copyright. All the institutions studied indicated that copyright clearance can be done to allow for the digitisation of materials that are under copyright. The University of Zimbabwe Library should, therefore, be prepared to seek clearance to digitise materials that are valuable but protected by copyright.

### **4.3.3 Guidelines on technologies**

The empirical study has shown that one type of scanner is not sufficient to perform digitisation in a library. The University of Zimbabwe should, therefore, ensure that money is available to purchase different type of scanners for scanning different types of materials or to then outsource a selected type of digitisation (i.e. books or maps). Flatbed scanners will have to be purchased for scanning documents and images; book scanners for scanning books and film scanners for scanning films. It is advisable that the institution chooses equipment that is being used by other institutions that are digitising so that it becomes easy to learn from those institutions about the experiences of using the equipment. It is not enough to rely on information from vendors of the equipment about the performance of the equipment. It is common that vendors give institutions information that will only serve to convince the institutions to buy. Learning from other institutions will highlight potential problems which might be associated with certain types of scanners.

Besides the scanners, there is need for the institution to buy a server for storing the digitised materials for archival purposes. The institution would also need to buy a second server to make provision for backups. Two servers with the capacity of storing up to 100 tera bytes would suffice for a start. One of the servers will have to be used for archiving the digitised materials while the other server will be used for back-up purposes. Bigger servers will need to be purchased as the number of digital materials that are added to the servers continues to increase.

Most of the software for the digitisation exercise comes with the equipment but there is need for the University of Zimbabwe Library to purchase some additional software that is required to complete the digitisation process. Software like ABBY FineReader will be required to convert scanned documents into readable text (Optical Character Recognition) and Adobe Photoshop will be required to edit some images and set the resolution to optimum levels for providing access to the digitised files on the web.

### **4.3.4 Guidelines on costs and sustainability**

This research has shown that digitisation is an expensive venture. Institutions planning to implement digitisation initiatives should be prepared to meet the expenses associated with digitisation. The major costs that the University of Zimbabwe Library should expect to meet are the costs of buying equipment and paying staff members' salaries. The institution should,

therefore, commit financial resources to meet these costs if the digitisation initiative has to be a success.

As was pointed out in the data analysis section, institutions can write grant application to raise funds for sustaining the digitisation exercise. The University of Zimbabwe can apply for grants such as the British Library Endangered Archives Programme (British Library 2012) where the British Library supports institutions in developing countries to digitise historical and cultural heritage materials and Elsevier Foundation's grants for innovative libraries in developing countries (Elsevier Foundation 2012) in which they support libraries embarking on innovative projects. Such grants may be used to fund digitisation at the University of Zimbabwe Library if the applications for the grants are successful.

The empirical study has also shown that digitisation can have the capacity to generate revenue. The University of Zimbabwe Library can, therefore, allow other institutions that intend to digitise to bring their materials for digitisation for a fee. The money generated from the digitisation of materials for external clients can be used to maintain the digitisation equipment so that the digitisation remains a sustainable programme. This, however, has implications for staff time and wear and tear of the equipment so the move to generate funds through digitising for other institutions should be well calculated to avoid losses over time.

There is hope that digitisation can be a sustainable method of preserving physical materials in the library from rapidly decaying due to over use. There is need for the institution's top management to commit themselves to supporting the digitisation at the institution by investing financial resources to cater for the costs of buying the equipment, maintaining the equipment and for paying the digitisation workforce. The institution need to plan strategically for digitisation and have a policy in place to support the digitisation. The policy has to be adopted by every stakeholder at the institution to guarantee sustainability of the digitisation.

Upgrading of equipment and software for digitisation has also been seen as one of the ways in which the digitisation of materials can be sustained and making sure that the digital objects can be preserved for long periods of time. Upgrading technologies, updating software and the use of open formats and open standards at the University of Zimbabwe Library can help in preserving digital objects for longer periods of time. The empirical study has shown that although the long-term preservation of digital materials is difficult, institutions which are totally committed to sustaining digitisation can be able to use digitisation as a sustainable method of preservation. The University of Zimbabwe Library should, therefore, try to embark on digitisation as a method of preservation and make it sustainable.



#### **4.3.5 Intellectual Property Issues**

As pointed out during the data analysis, copyright clearance has to be sought for all materials that are not the intellectual property of the digitising institution. A template letter has been drafted to help the University of Zimbabwe Library seek clearance for digitising materials under copyright (see Appendix B). A staff member responsible for seeking copyright clearance has to be appointed to ensure that the copyright clearance is done without any delays.

Digitising out-of-copyright materials (or materials in the public domain) and materials whose copyright belongs to the University of Zimbabwe can help the institution avoid violating copyright. The institution can choose to digitise those materials that are in the public domain and materials that the institutions own copyright while clearance is being sought to digitise those materials that are still protected under copyright. This would mean that the digitisation will always continue to run even if some of the materials may take some time before permission to digitise them has been obtained.

#### **4.3.6 Steps in digitisation**

An analysis of the steps that the institutions that were studied follow in their digitisation initiatives and the steps as found in the literature study (Section 2.5, Chapter 2) yields nine distinct steps. The steps outlined below can be followed by the University of Zimbabwe in their digitisation:

- **Selection** – Selection involves the assessment of the value of materials to be digitised. This is done so that only materials that are of value can be digitised. Materials of historical significance and those that describe local history and written by local authors are certainly of value to the University of Zimbabwe Library. Such materials would need to be prioritised for digitisation.
- **Copyright checks** – This involves determining the copyright status of materials for digitisation. If the materials are out of copyright, or if the copyright of the materials belongs to the University of Zimbabwe then materials can be digitised. If the

materials are under copyright, permission to digitise the materials can be sought. The copyright clearance template (Appendix B) should be used to request for permission to digitise.

- **Document preparation** – Document preparation involves preparing the materials for the scanner. If documents are stapled, the staples have to be removed so that the documents can be easily fed into a flatbed or feeder scanner. Where it is cost effective to remove the binding from books, bound books can be unbound to make them easier to scan.
- **Scanning** – This involves feeding the material into the scanner, setting the scan resolution and image capture. Scanning produces a digital copy of the document.
- **Metadata creation** – Creation of metadata involves describing the digital object so that it can be easily identified. The metadata allows the institution to have intellectual control over the digitised materials as well as make it easy for users to retrieve the digitised materials.
- **Quality control** – Quality control involves verifying that the scanning and the metadata creation have been done according to standards. The University of Zimbabwe Library should appoint a staff member dedicated for quality control to ensure that everything is done according to standard. If the scanning and the metadata are not in conformity with the standards set for the institution, it has to be corrected.
- **Upload to archival server** – After the quality has been checked, it is now necessary to upload the scanned material onto the archival server. The material will have to be preserved on the server. Copies of the digitised materials will need to be deposited in the back-up server.
- **Optimisation for web publishing** – Both the literature study (section 2.5, Chapter 2) and the empirical study showed that the standards for archiving and for providing access to digitised objects are different. For materials that have to be accessed on the Internet, there is need to convert the digital files into file formats that are suitable for access through the Internet. Optical Character Recognition software has to be used to allow text documents to be searchable; TIFF files will need to be converted to

PDF to reduce file size and in the case of images, TIFF files will have to be converted to JPEG format. All this allows for easy and quick downloading from the Internet.

- **Publish on web** – The ultimate goal of digitisation is to provide easy access to the digitised files on the Internet. The final step, therefore, has to be to make the files accessible on the Internet. The digital objects can be published through the Institutional Repository. This will give registered users of the University of Zimbabwe Library the platform to access the digitised materials.

Following the steps outlined above will ensure that the University of Zimbabwe Library engages in a serious digitisation exercise. The digitisation process would be a comprehensive one and other institutions would be able to learn from the institution.

#### **4.3.7 Challenges in digitisation**

The University of Zimbabwe Library has to be ready for the challenges that have been identified in this research as hampering digitisation projects. Being aware and being prepared to tackle the challenges of digitisation will assist the institution in its digitisation initiatives. This guide is, therefore, going to suggest measures that can be put in place to counter or mitigate the challenges that may be faced in digitisation at the University of Zimbabwe Library.

Human resources pose arguably the greatest challenge in digitisation as seen in this research where four out of the five institutions that were studied pointed out that human resources have proved to be one of the challenges that affect their digitisation initiatives. Attracting and retaining human resources for the digitisation initiatives is critical. The institution needs to have the strategies in place to ensure that they attract and retain staff for the digitisation. One of the strategies that the University of Zimbabwe can put in place to attract and retain staff is to continuously develop their staff. The staff will feel motivated if they continuously get opportunities to advance their careers.

Adequate human resources have to be in place at each stage in the digitisation process to ensure that every step that has to be done will be done properly and thoroughly. The staff should also be motivated so that they perform their duties with diligence and integrity. This

would ensure that the digitisation remains sustainable. It would be very difficult to adhere to best standards and practises if the staff members are not motivated or inadequate.

Besides having adequate staff that is motivated to do the digitisation, the staff members must have the requisite skills to be able to complete the digitisation process. It is therefore recommended that selected staff members are sent for training in institutions that are already digitising. The institutions that have already started digitisation have experiences and skills that they can share with other institutions. Institution D may be a viable option as the pointed out that they provide training to other institutions that need to learn the digitisation skills.

Another challenge that is inevitable in digitisation is keeping up with technology. It is recommended that the University of Zimbabwe Library should choose technologies that are already in use at other institutions. This helps in sharing experiences, troubleshooting problems and learning from each other. A budget has to be set aside to ensure that there are financial resources to upgrade technologies that need upgrading.

The study has shown that the cost of equipment for digitisation is expensive. The University of Zimbabwe Library has to be prepared to meet the costs of buying and maintaining the necessary equipment for digitising or the effort should be aborted. Although this is going to be expensive, the benefits to be derived from digitisation are immense so it is worthwhile for the institution to invest in the digitisation equipment. If the costs of buying and maintaining equipment prove to be beyond the budget of the institution, the digitisation can be outsourced from other institutions which already have the equipment. The empirical study has shown that outsourcing digitisation from other institutions is an option that other institutions who cannot afford to buy their own equipment are using. This is done because it is cheaper to outsource the digitisation than to buy and maintain the expensive equipment.

Another challenge that the University of Zimbabwe Library has to be ready for is getting approval for the sustainable digitisation of materials. The institution has to seek the approval of all the major stakeholders before embarking on the digitisation. It has been emphasised from the empirical study that it is very important to get the approval and commitment from top management for the digitisation to be a success. Getting the approval and commitment means that top management will commit to sustain the digitisation. One way of getting the approval is through involving top management in the initial planning for digitisation and show them the benefits of digitisation. If the approval and commitment is not there, there is a risk

that the digitisation may not be sustainable in the future. The success of the digitisation exercise hinges on the continued financial support from top management.

#### **4.4. Conclusion**

This chapter presented an analysis of the data that was gathered through the use of semi-structured interviews with digitisation managers from five institutions. The analysis provided a platform for guidelines for digitisation for the University of Zimbabwe Library. The guidelines presented were based on the researcher's interpretation of the research findings from the empirical research and the findings from the literature review conducted in Chapter 2. The next chapter, will summarise the research findings, provide recommendations and give conclusions regarding digitisation as a method of preserving information.

## **Chapter 5**

### **5. Conclusions and Recommendations**

#### **5.1. Introduction**

Chapter 1 of this research introduced the research and gave background information about the research. Chapter 2 focused on reviewing the available literature on digitisation, while Chapter 3 gave a detailed report on the methodology that was used to collect and analyse data for the research. This was followed by Chapter 4 which gave an analysis and interpretation of the data as well as present guidelines for digitisation at the University of Zimbabwe Library based on the findings from the literature study (Chapter 2) as well as the empirical study. The purpose of this final chapter is to give conclusions and recommendations based on the findings from a study of five institutions which are digitising their collections.

#### **5.2. Conclusions and Recommendations**

The conclusions and recommendations given here address the problems that this study sought to solve. Thus, the conclusions and recommendations fall under the following topics as defined by the research questions (see section 1.2.1 in Chapter 1): best practices and standards, selection of materials, technologies for digitisation, costs of digitisation, steps in digitisation and intellectual property issues.

##### **5.2.1 Best practices and standards**

There are quite a number of standards that are used in digitisation as has been seen in Chapter 4. The standards that are widely used for archiving digitised materials are:

- Archival files: resolution of 600 dpi and archived in the original TIFF format – a compromise between quality and size of the files is made. These are seen as the

optimal standards that will satisfy quality requirements and the need to save on disk space.

- Access Files: resolution of 200 dpi and access is provided in PDF format or JPEG format – a compromise between the quality of the display and the time required to download the file is made. The smaller the size of the digitised object, the lesser the time required to download it but also it means the quality is compromised to some degree.

Despite strong arguments for MODS and METS (See section 2.2.3 in Chapter 2) as metadata schema that are flexible for capturing descriptive, administrative and structural metadata, Dublin Core has proved to be the most used metadata schema in most academic institutions. Dublin Core is preferred because of its easy of use, flexibility and interoperability with MARC21 (NISO 2004). Many academic institutions use it so it is recommended that the University of Zimbabwe Library should use it as their metadata standard for digitised materials.

### **5.2.2 Selection of materials**

This research has demonstrated that institutions cannot digitise everything in their collection no matter how much financial resources they have. Digitisation is an expensive venture. This means there has to be a serious process of selection of materials for digitisation. It is imperative that an institution digitises collections that are of value to their clientele. Different criteria have to be applied in the selection of materials for digitisation. These criteria include but are not limited to the value of the collection, the cost of digitising the collection and the purpose for digitising the collection.

### **5.2.3 Technologies for digitisation**

The technologies for digitisation are varied. This research has shown that one type of scanner is not sufficient to digitise the different types of materials that libraries have. Flatbed scanners will be required to digitise unbound materials while book scanners which use mostly camera technology will be required to digitise bound books and oversized materials and film scanners will be required to digitise microfilm and slides. The University of Zimbabwe Library should purchase at least a flatbed scanner, a book scanner and a film

scanner for digitising materials in its special collections. The brand of the scanners that the institution should choose would be influenced by the affordability of the scanners in relation to the budget available.

#### **5.2.4 Cost of digitisation**

Digitisation does not come cheap. Equipment and human resources contribute to the high costs in digitisation. Institutions that intend to digitise their collections must be prepared to fund their digitisation initiatives. It has, however, been found out that institutions are able to apply for grants from donor institutions which support digitisation initiatives. The grants can be used to support the digitisation activities but institutions will have to find sustainable ways of running their digitisation projects. One way of doing that would be to digitise materials for other institutions for a fee. The fees received from digitising materials for other institutions could be used to sustain the digitisation activities.

The study also showed that institutions that cannot afford to buy equipment and pay a full complement of digitisation staff can outsource digitisation. The University of Zimbabwe Library could take the option of outsourcing if financial resources are not sufficient to buy all the equipment and pay salaries for digitisation staff. This could prove to be viable because some institutions that were studied in this research have managed to do their digitisation through outsourcing.

#### **5.2.5 Intellectual property issues**

The research also found out that intellectual property issues cannot be ignored when digitising collections. The issue of copyright is one that has to be sorted out if digitisation is going to be successful. It is critical that copyright clearance be sought for materials that are protected by copyright before any digitisation can begin. This helps the institution to avoid any legal battles that may arise because of digitising materials without permission from the copyright owners.

Besides having to deal with copyright issues in digitisation, it is clear that intellectual property issues cannot hinder digitisation initiatives if the necessary measures are taken to avoid violating copyright. The measures that can be employed to avoid violating copyright include digitising materials whose copyright belong to the digitising institution, digitising materials



that are in the public domain and requesting for permission to digitise works from the publishers or the authors of the works.

### **5.2.6 Steps in digitisation**

The literature study and empirical study that was carried out identified the steps that are followed in digitisation. Although these steps vary from institution to institution, the steps were analysed and combined to come up with a solid workflow that can be used by any organisation. Systematically following the steps that have been recommended for the University of Zimbabwe Library would enable institutions to run successful digitisation projects.

### **5.2.7 Other issues for consideration**

The use of a DMS, for managing the digitisation workflow is recommended. Although none of the institutions studied is using a DMS, Institution A, which is one of the first institutions to embark on digitisation, highly recommended the use of DMS for managing digitisation workflow effectively. This would ensure that all the steps in the digitisation process are followed (see section 4.2.8).

Given that there are so many steps that are involved in the digitisation process, there is need to have adequate manpower to do the digitisation process. As has been seen from the empirical study, the need to have capable human resources at all the stages in the digitisation cycle cannot be overemphasised. The University of Zimbabwe is therefore recommended to scout for human resources who are capable of completing the digitisation exercise. Although there are no formal courses for digitisation in higher education, Institution D pointed out that they are in a position to be able to train staff from other institutions who intend to embark on digitisation.

It has become clear from this research that there is need to be aware of the challenges that affect the digitisation of materials in academic libraries. Being aware of these problems before embarking on digitisation helps institutions to be prepared to tackle these challenges when they arise. For example the research found out that keeping up with technological changes is one of the challenges that affect digitisation. An institution planning to embark on

digitisation can, therefore, come up with strategies on managing the changes in technology so that the changes will not have a greater negative effect on the digitisation.

As has been seen from the empirical research, it is important for institutions that intend to embark on serious digitisation to have digitisation policies. This is important because of the need to get buy-in from top management for the digitisation initiatives. The University of Zimbabwe Library needs to have the digitisation policy in place before any digitisation begins. The policy will help to justify the actions of digitisation staff and will also mandate top management to support sustainable digitisation at the institution.

### **5.3. Future and Further studies**

The rate at which technology is changing is alarming. It would be interesting to do a study after five years to observe which technologies would be in use for digitisation. The standards for digitisation are also likely to change in future in line with the changes that would have taken place in technology. It would also be interesting to map how the standards for digitisation have changed over the years.

The respondents in this research have given so much hope that digitisation can be used as a sustainable method of preservation of information. It would be important for research to be done in institutions that would have tried to use digitisation as a sustainable method of preservation. It would be useful to measure the success that those institutions would have registered.

### **5.4 Final remarks**

This research has come up with guidelines for digitisation at the University of Zimbabwe Library which if followed could result in the implementation of a successful digitisation initiative at the institutions. It is, however, important to ensure that digitisation staff go for training not only for them to be able to understand these guidelines but also to be able to execute the digitisation effectively and efficiently. Other academic institutions in Zimbabwe could also benefit from these findings. The institutions could adopt the guidelines with minor alterations.

## References

Babbie, E. & Mouton, J. 2001. *The practice of social research*. Oxford University Press, Cape Town.

Bailey, C. W., 2011. *Digital Curation and preservation bibliography 2010*. Available at: <http://digital-scholarship.org/dcpb/dcpb2010.htm> [Accessed February 16, 2012].

Bailey, K. D., 1982. *Methods of Social Research, 2nd ed.* The Free Press, London.

Bailey-Hainer, B. & Urban, R., 2004. The Colorado digitization program: a collaboration success story. *Library Hi Tech*, 22(3), pp.254-262. Available at: <http://www.emeraldinsight.com/10.1108/07378830410560044> [Accessed February 15, 2012].

Balas, J. I., 2011. Digitisation on a mission Part II: Walking the walk. *Computers in Libraries*. Available at: <http://www.infotoday.com> [Accessed February 19, 2012].

Ballard, T. & Donald, E., 2007. A digitization and multimedia project at Quinnipiac University, Hamden, Connecticut, USA. *New Library World*, 108(9/10), pp.445-452. Available at: <http://www.emeraldinsight.com/10.1108/03074800710823971> [Accessed December 6, 2011].

Berger, M., 1999. Digitization for preservation and access : a case study. *Library Hi Tech*, pp.146-151.

Bia, A., Muñoz, R. & Gómez, J., 2011. DiCoMo: the digitization cost model. *International Journal on Digital Libraries*, 11(2), pp.141-153. Available at: <http://www.springerlink.com/index/10.1007/s00799-011-0073-9> [Accessed February 15, 2012].

Book Drive, 2012. Available at: <http://diy.atiz.com/> [Accessed April 20, 2012].

Brindley, L., 2012. Investing in digitisation. *Information Today*, January 2012, pp.14-16. Available at: <http://www.infotoday.com> [Accessed February 19, 2012].

British Library, 2012. *British Library Endangered Archives Programme*. Available at: <http://eap.bl.uk/> [Accessed July 16, 2012].

B&H Photo video, 2012. *Film scanners*. Available at: <http://www.bhphotovideo.com/c/buy/Film-Scanners/ci/1151/N/4289285573> [Accessed April 20, 2012].

Chapman, S. & Kenney, A. R., 1996. Digital Conversion of Research Library Materials. *D-Lib Magazine*. Available at: <http://dlib.org/dlib/october96/cornell/10chapman.html> [Accessed April 20, 2012].

Computer Supplies unlimited. 2012. *HP Scanjet N9120 Document Flatbed Scanner*. Available at: <http://www.computersuppliesultd.com/Peripherals/Scanners/L2683A.html> [Accessed April 20, 2012].

- Creswell, J. W., 2009. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*, 3rd ed. Sage Publications, London.
- Dames, K., 2005. Copyright clearances: fair use, action and apathy, *ONLINE magazine*, 29(5), pp. 32-34. Available at: <http://www.onlinemag.net> [Accessed April 30, 2012].
- Denscombe, M. 2003. *The Good Research Guide: For small-scale social research projects*, 3rd ed. Open University Press, New York.
- Eakin, L., Friedlander, A. & Schonfeld, R., 2006. *A Selective Literature Review on Digital Preservation Sustainability*. Available at: [http://brtf.sdsc.edu/biblio/Cost\\_Literature\\_Review.pdf](http://brtf.sdsc.edu/biblio/Cost_Literature_Review.pdf) [Accessed December 6, 2011].
- Eden, B., 2001. Getting started with library digitization projects : funding your first ( and subsequent ) digital projects. *The Bottom Line: Managing Library Finances*, 14(2), pp.53-55.
- Elsevier Foundation, 2012. *Innovative Libraries*. Available at: <http://www.elsevierfoundation.org/innovative-libraries/> [Accessed July 16, 2012].
- Evens, T. & Hauttekeete, L., 2011. Challenges of digital preservation for cultural heritage institutions. *Journal of Librarianship and Information Science*, 43(3), pp.157-165. Available at: <http://lis.sagepub.com/cgi/doi/10.1177/0961000611410585> [Accessed November 20, 2011].
- Gartner, R., 2008. Metadata for digital libraries: state of the art and future directions. *JISC Technology and Standards Watch*, April. 2008. Available at: [http://www.jisc.ac.uk/media/documents/techwatch/tsw\\_0801pdf.pdf](http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf) [Accessed April 30, 2012].
- Gbaje, E.S., 2011. Digital preservation strategies: A case study of Nigerian national information centres. *IFLA Journal*, 37(3), pp.218-227. Available at: <http://ifl.sagepub.com/cgi/doi/10.1177/0340035211418727> [Accessed March 3, 2012].
- Gorman, G. E. & Clayton, P., 1997. *Qualitative Research for the Information Professional: a practical handbook*. Library Association Publishing, London.
- Gravetter, F. J. & Forzano, L. B., 2009. *Research Methods for the Behavioral Sciences*, 3rd ed. Wadsworth Cengage Learning, Belmont.
- Green, A., 2010. Big digitisation: Origins, progress and prospects. *International Journal of Humanities and Arts Computing*, 4(1-2), pp.55-66. Available at: <http://www.eupublishing.com/doi/abs/10.3366/ijhac.2011.0007> [Accessed December 22, 2011].
- Groenewald, R. & Breytenbach, A., 2011. The use of metadata and preservation methods for continuous access to digital data, *Electronic Library*, 29(2), pp.236 - 248. Available at: <http://www.emeraldinsight.com/journals.htm?articleid=1917480> [Accessed April 2, 2012].
- Groenewald, R. 2010. 'Collection Development', *In National Research Foundation, 2010. Managing digital collections: a collaborative initiative on the South African Framework*. Available at: <http://digi.nrf.ac.za/publ/Managing%20Digital%20Collections.pdf> [Accessed July 02, 2012].

Groenewald, R. & Klapwijk, W. 2010. 'Objects', In *National Research Foundation, 2010. Managing digital collections: a collaborative initiative on the South African Framework*. Available at: <http://digi.nrf.ac.za/publ/Managing%20Digital%20Collections.pdf> [Accessed July 02, 2012].

Guenther, R & McCallum, S., 2003. New Metadata Standards for Digital Resources: MODS and METS. *Bulletin of the American Society for Information Science and Technology*, December/January 2003. Available at: <http://onlinelibrary.wiley.com/doi/10.1002/bult.268/pdf> [Accessed April 26, 2012].

Hammond, M. & Davies, C., 2009. *Understanding the costs of digitization: detailed report*. Available at: <http://www.jisc.ac.uk/media/documents/programmes/digitisation/digitisation-costs-full.pdf> [Accessed April 30, 2012].

Hasenay, D. & Krtalic, M., 2010. Preservation of newspapers: Theoretical approaches and practical achievements. *Journal of Librarianship and Information Science*, 42(4), pp.245-255. Available at: <http://lis.sagepub.com/cgi/doi/10.1177/0961000610380818> [Accessed March 3, 2012].

Hazen, D., Horrell, J. & Merrill-Oldham, J., 1998. *Selecting Research Collections for Digitization*. Washington, DC: Council on Library and Information Resources. Available at: <http://www.clir.org/pubs/reports/hazen/pub74.html> [Accessed April 28, 2012].

Hockx-Yu, H., 2006. Digital preservation in the context of institutional repositories. *Program: electronic library and information systems*, 40(3), pp. 232-243. Available at: DOI 10.1108/00330330610681312 [Accessed January 28, 2012]

Hodge, G., 2004. *Preservation of and Permanent Access to Electronic Information Resources*. Available at: <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA432687> [Accessed February 21, 2012].

Hsieh, H-F. & Shannon, E. S., 2005. Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, 15(9), pp. 1277-1288.

Hughes, L., 2004. *Digitizing Collections: Strategic Issues for the Information Manager*, Facet Publishing, London.

Huwe, T. K., 2011. Involving the community in digitisation. *Computers in Libraries*, February 2011, pp. 21-24. Available at: <http://www.infotoday.com> [Accessed February 19, 2012].

i2S Corporate, 2012. Available at: [http://www.i2s-bookscanner.com/produits.asp?gamme=1024&sX\\_Menu\\_selectedID=left\\_1024\\_GEN](http://www.i2s-bookscanner.com/produits.asp?gamme=1024&sX_Menu_selectedID=left_1024_GEN) [Accessed April 28, 2012].

International Federation of Film Archives, 2012. *Choosing a film scanner*. Available at: <http://www.fiafnet.org/commissions/TC%20docs/Choosing%20a%20Scanner%20v1%201.pdf> [Accessed October 6, 2012].

ITU, 2012. *Key statistical highlights: ITU data release*, June 2012 Available at: <http://www.itu.int/ITU->

[D/ict/statistics/material/pdf/2011%20Statistical%20highlights\\_June\\_2012.pdf](#) [Accessed October 26, 2012].

Iwhiwhu, B.E. & Eyekpegba, E.O., 2009. Digitization of Nigerian university libraries: From technology challenge to effective information delivery. *The Electronic Library*, 27(3), pp.529-536. Available at: <http://www.emeraldinsight.com/10.1108/02640470910966943> [Accessed December 6, 2011].

Jha, N.K., 2008. *Research Methodology*. Abhisek Publications, Chandigarh. Available at: <http://site.ebrary.com/lib/uzlib/Doc?id=10416498&ppg=46> [Accessed July 20, 2012].

JISC., 2009. *JISC Project Report : Digitisation Programme*. Preservation Study, April 2009.

Jones, R.A., 2005. Empowerment for digitization: lessons learned from The Making of Modern Michigan. *Library Hi Tech*, 23(2), pp.205-219. Available at: <http://www.emeraldinsight.com/10.1108/07378830510605160> [Accessed February 15, 2012].

Kumar, K., 1999. *Research Methods in Library and Information Science, 2<sup>nd</sup> ed.* Har-Anand Publications, Delhi.

Kupfer, S., 2010. A rising tide of digitization - The Ohio Memory Project. *Computers in Libraries*, Available at: <http://www.infotoday.com> [Accessed February 16, 2012].

Lampert, C. & Vaughan, J. 2009. Success factors and strategic planning: Rebuilding an academic library digitization program. *Information Technology and Libraries*, pp. 116-137

Lee, S., 2001. *Digital Imaging: A Practical Handbook*. Neal-Schuman, New York, NY.

Liu, Y.Q., 2004. Best practices, standards and techniques for digitizing library materials: a snapshot of library digitization practices in the USA. *Online Information Review*, 28(5), pp.338-345. Available at: <http://www.emeraldinsight.com/10.1108/14684520410564262> [Accessed February 15, 2012].

Lopatin, L., 2006. Library digitization projects, issues and guidelines: A survey of the literature. *Library Hi Tech*, 24(2), pp.273-289. Available at: <http://www.emeraldinsight.com/10.1108/07378830610669637> [Accessed September 16, 2011].

Messmer, G., 2011. *Digitization Workflow of the Bavarian State Library*. Available at: [http://www.digitale-sammlungen.de/~mdz/mdz/content/service/docs/Messmer\\_Digitization\\_Workflow.pdf](http://www.digitale-sammlungen.de/~mdz/mdz/content/service/docs/Messmer_Digitization_Workflow.pdf) [Accessed February 16, 2012].

Minow, M. & Smith, K., 2011. Keeping Library Digitization Legal. *American Libraries Magazine*, Available at: <http://americanlibrariesmagazine.org> [Accessed February 16, 2012].

Munir, K.A. & Phillips, N., 2002. *The concept of industry and the case of radical technological change*. *Technology Management*, 13, pp.279-297.

Nicholson, D. 2010. 'Copyright and related matters', *In National Research Foundation, 2010. Managing digital collections: a collaborative initiative on the South African Framework*.

Available at: <http://digi.nrf.ac.za/publ/Managing%20Digital%20Collections.pdf> [Accessed July 02, 2012].

NISO, 2004. *Understanding Metadata*. Available at: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>, [Accessed December 6, 2011].

Northam, A., 2009. *Basic digitization – scanning toolkit*. Available at: <http://www.slideshare.net/heirloomproject/basic-digitization-scanning-toolkit2-voice> [Accesses April 28, 2012].

Novara, E. A., 2010. Digitization and researcher demand: Digital imaging workflows at the University of Maryland Libraries. *OCLC Systems & Services*, 26(3), pp.166-176. Available at: <http://www.emeraldinsight.com/10.1108/10650751011073616> [Accessed August 26, 2011].

OCLC/RLG Working Group on Preservation Metadata, 2001. *Preservation Metadata for Digital Objects: A Review of the State of the Art. A White Paper by the OCLC/RLG Working Group on Preservation Metadata*. Available at: [http://www.oclc.org/research/activities/past/orprojects/pmwg/presmeta\\_wp.pdf](http://www.oclc.org/research/activities/past/orprojects/pmwg/presmeta_wp.pdf) [Accessed April 29, 2012].

Peterson, K. A., 2005. *What to Look for in a Scanner: Tip Sheet for Digitizing Pictorial Materials in Cultural Institutions*. Available at: <http://www.loc.gov/rr/print/tp/LookForAScanner.pdf> [Accessed April 28, 2012].

Petty, N. J.; Thomson, O. P. & Stew, G., 2012. Ready for a paradigm shift? Part 2: Introducing qualitative research methodologies and methods. *Manual Therapy*, 17(37)

Pienaar, H., 2001. *Design of a web portal for academics (Afrikaans)*. M.Ed Computer-Assisted Education, University of Pretoria, Pretoria Available at: <http://upetd.up.ac.za/thesis/available/etd-03202003-161802/> [Accessed October 15, 2012]

Puglia, S.; Reed, J. & Rhodes, E., 2004. *NARA Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images*. Available at: <http://www.archives.gov/preservation/technical/guidelines.pdf> [Accessed April 29, 2012].

Punch, K. F., 2000. *Developing effective research proposals*. Sage Publications, London.

Republic of Zimbabwe, 2000. *Copyright and Neighbouring Rights Act: Chapter 26:05, Act 11/2000, 22/2001 (s. 4), 32/2004*.

Rieger, O. Y., 2008. *Preservation in the Age of Large-Scale Digitization: A White Paper*. Available at: <http://www.clir.org/pubs/abstract/pub141abst.html>. [Accessed January 03, 2012]

Riley, J., 2004. *IU DLP Use of Digital Imaging Standards and Best Practices*. Available at: <http://www.dlib.indiana.edu> [Accessed April 28, 2012].

Robson, C., 2011. *Real world research. 3rd ed*. Wiley, Chichester.

Sandelowski, M., 2000. Focus on Research Methods: Combining Qualitative and Quantitative Sampling, Data Collection, and Analysis Techniques in Mixed-Method Studies. *Research in Nursing & Health*, 23(3), pp. 246-255.

Tanner, S., 2006. *Handbook on Cost Reduction in Digitisation*. Available at: [http://www.minervaeurope.org/publications/CostReductioninDigitisation\\_v1\\_0610.pdf](http://www.minervaeurope.org/publications/CostReductioninDigitisation_v1_0610.pdf) [Accessed April 30, 2012].

United Nations ARMS, 2006. *Guideline on Records Digitisation*. Available at: <http://archives.un.org/unarms/en/unrecordsmgmt/unrecordsresources/GUIDELINE%20ON%20RECORDS%20DIGITISATION.htm> [Accessed April 27, 2012].

University of KwaZulu-Natal, 2010. *Understanding research, 3<sup>rd</sup> ed.*, School of Education and Development, KwaZulu-Natal.

University of Zimbabwe, 2012. *University of Zimbabwe Website*. Available at: <http://www.uz.ac.zw> [Accessed June 12, 2012].

Van der Merwe, 2011. Digitization and migration activities at the CSIR , RSA : a case study. *4th African Conference for Digital Scholarship and Curation*. Available at: <http://www.nedicc.ac.za/Conference/Upload%20Papaers/van%20der%20Merwe-Digital%20curation.pdf> [Accessed February 19, 2012].

Vogt-O'Connor, D. (2000), "Selection of materials for scanning", in Sitts, M. (Ed.), *Handbook for Digital Projects: A Management Tool for Preservation and Access*, Northeast Document Conservation Center, Andover, MA.

Van Deventer, MJ. 2009. Selective Digitisation of Information: The CSIR ' s strategy for a sustainable effort. In *1st International Conference on African Digital Libraries and Archives*. Addis Ababa: 1-3 July 2009.

Vrana, R., 2011. Organizational aspects of digitization projects in the Croatian public libraries. *Library Management*, 32(8/9), pp.589-598. Available at: <http://www.emeraldinsight.com/10.1108/01435121111187941> [Accessed December 6, 2011].

Vrana, R., 2010. Public libraries in Croatia and the digitization challenge. *Library Review*, 59(5), pp.325-340. Available at: <http://www.emeraldinsight.com/10.1108/00242531011047028> [Accessed December 6, 2011].

Wentzel, L., 2006. Scanning for Digitization Projects. *Library Hi Tech News*, 23(4), pp.11-13. Available at: <http://www.emeraldinsight.com/10.1108/07419050610674712> [Accessed December 29, 2011].

Young, F. W., 1981. Quantitative analysis of qualitative data. *Pschometrica*, 46(4), pp.357 – 388.



## **Appendix A – Interview schedule**

### **Interview schedule for extracting/ determining digitisation guidelines from South African institutions**

#### **1 Text of an e-mail to be sent to prospective interviewees**

My name is Lovemore Kusekwa, M.IT student at the University of Pretoria. One of the requirements of the Masters Degree programme is that students should produce a mini-dissertation in partial fulfilment of the degree. My dissertation focuses on producing guidelines for digitisation at the University of Zimbabwe where I am employed. As you are responsible for digitisation at your institution you were identified as a candidate who would be able to assist and inform me regarding such guidelines.

I would like to gain your opinion regarding the following:

- the digitisation standards and best practices at your institution,
- selection criteria of materials for digitisation,
- the technologies that you are using in digitisation,
- costs of digitisation and sustainability issues,
- challenges in digitisation,
- intellectual property rights, and
- steps that you follow in the digitisation process.

I intend using information collected from a number of interviewees to develop guidelines for digitisation for my institution. I also hope that the guidelines could be useful to other institutions that are planning to digitise their collections.

The interview should take about one to one and a half hour. Are you available to respond to some questions at this time? If so: could you please identify a suitable timeslot on any of the days between 30 July 2012 and 08 August-2012 when I shall be in South Africa to conduct the interviews.

## **2.1 Introduction**

Thank you for making time available to discuss issues related to digitisation with me. As I indicated in my e-mail I would like to discuss matters related to the digitisation standards and best practices at your institution, selection criteria of materials for digitisation, the technologies that you are using in digitisation, costs of digitisation and sustainability issues, challenges in digitisation, intellectual property rights, and the steps that you follow in the digitisation process. I will also provide you with an opportunity to add any additional aspects that you feel I should take note of.

This interview is going to be recorded to assist with the coding of the information. The information gathered through this interview will be handled confidentially. Are you comfortable with responding to questions on the record?

## **2.2 Questions**

### **A. Digitisation standards and best practices**

1. Which standards do you use in digitisation?
2. Are you using the same standard(s) for giving access to digitised materials as those you use to archive the digitised materials?
3. Please explain why you chose the standard(s) that you have mentioned.
4. Which metadata standard(s) do you use for your digitised materials?
5. Kindly explain why you chose the metadata standard(s) that you are using.
6. Which file formats do you use for archiving and for giving access to digitised materials?

### **B. Selection criteria**

1. Which criteria do you use to select materials for digitisation?
2. Please explain why each criterion you have mentioned is important

### **C. Technologies for digitisation**

1. Which technologies are you using for digitisation?
2. Please explain in full all the equipment that you are using for digitisation.
3. Which factors have influenced the choice of technologies that you are using?

### **D. Costs of digitisation and sustainability issues**

1. What are the costs involved in digitisation?
2. How are the expenses/ costs funded?
3. Given the complexities and expenses, do you think digitisation can be used as a sustainable method of preservation?

E. Intellectual property rights issues

1. What are the intellectual property rights issues that you have had to contend with in your digitisation programme?
2. How do intellectual property laws in your country help or constrain you when it comes to digitisation?
3. How do you ensure that you do not violate copyright when doing digitisation?

F. Steps in digitisation

1. Do you make use of a project-based approach to your digitisation? If not, what approach do you use? Please explain the reasons for your approach.
2. What steps do you include in your digitisation process?
3. Why is each of the steps that you have mentioned necessary?

G. Challenges in digitisation

1. What major challenges are you facing in digitising your collections?
2. How do you think these challenges can be overcome?

H. Other issues that should be considered

1. Are there any other issues you would like us to discuss at this time? Is there anything else you think would be helpful for me to know so that I can successfully come up with guidelines for digitisation in my institution?

### **2.3 Closing Remarks**

I appreciate the time you took for this interview.

Should I have any further questions relating to the subject that we have discussed today, would it be all right for me to contact you?

Would you like to see the analysed results from the interviews I conducted?

Thank you once again, I look forward to successfully using the information that you have shared with me today to come up with guidelines that can be useful to my institution and other institutions that are planning to digitise.

## Appendix B – Copyright Clearance Template

### University of Zimbabwe Library



### Copyright Clearance Template

---

DATE

Addressee

city

Country

Dear [Copyright holder],

On behalf of the University of Zimbabwe Library, I am writing to request for permission to digitise your work (s) listed below. We intend to digitise the works, retain the digital copy and make them available on the web to our registered users of the library. Our users are comprised of students, researchers, administrative staff and academic staff.

The work(s) for which the copyright permission is requested are:

[List the work(s) here]

The digitised work(s) will be used to support teaching, learning and research at our institution.

Please review and sign this letter to confirm that you are the rights-holder for this work, or have authority to grant permission to use the work, and that you grant permission to the University of Zimbabwe Library according to the terms outlined above.

---

Signature

If you are not the copyright holder, please refer us to the person that you believe is the copyright holder if you know them by indicating their name and contact details here:

.....  
.....  
.....  
.....  
.....

Thank you in advance for supporting us.

Sincerely,

Digitisation Manager  
University of Zimbabwe Library  
P.O. Box MP45, Mt. Pleasant, Harare, Zimbabwe  
+263 4 303 211  
[digitisation@uzlib.uz.ac.zw](mailto:digitisation@uzlib.uz.ac.zw)