

**EFFICACY OF DIGITALPRESERVATION PRACTICES IN INSTITUTIONAL  
REPOSITORIES (IRs) OF SELECTED PUBLIC UNIVERSITIES IN KENYA**

**BY**

**HELLEN MUKIRI NDEGWA**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE AWARD OF DOCTOR OF PHILOSOPHY  
DEGREE IN LIBRARY AND INFORMATION STUDIES, DEPARTMENT OF  
LIBRARY, RECORDS MANAGEMENT AND INFORMATION STUDIES,  
SCHOOL OF INFORMATION SCIENCES**

**MOI UNIVERSITY**

**ELDORET**

**2023**

**DECLARATION**

**DECLARATION BY THE CANDIDATE:**

This thesis is my original work and has not been presented for a degree in any other university. No part of this thesis may be reproduced without the prior written permission of the author and/or Moi University.

HELLEN M. NDEGWA

IS/DPHIL/01/12

Signed: ..... Date: .....

**DECLARATION BY THE SUPERVISORS:**

We confirm that the work reported in this thesis was carried out under our supervision and has been submitted with our approval as University supervisors.

Signed ..... Date: .....

Dr Emily Bosire

Department of Library, Records management and Information Studies

School of Information sciences

Moi University, Eldoret

Signed: ..... Date: .....

Dr. Damaris Odero

Department of Library, Records Management and Information Studies

School of Information Sciences

Moi University, Eldoret

## **DEDICATION**

I dedicate this thesis to God Almighty for the far that he has brought me and to my family as well all those who are on an academic journey irrespective of the level. Believe in your-self.

## **ACKNOWLEDGMENT**

First, I would like to express my sincere gratitude to my supervisors Dr. Emily Bosire and Dr. Damaris Odera for their unwavering support during this research. Your guidance has really been a blessing not only to my research but to my career.

Similarly, profound gratitude goes to my colleagues at Karatina University for their encouragement even when the going got tough. It has been a journey. You are true heroes.

Not to be forgotten are my research participants. Your enthusiasm made the hustle worth.

Finally, my sons Nick and Dan and other family members for giving me the challenge that you will graduate with a PHD before me if I didn't make the effort. At least you will have to come behind me now. May God see you through your academic journeys.

## TABLE OF CONTENTS

DECLARATION .....	ii
DECLARATION BY THE CANDIDATE: .....	ii
DEDICATION .....	iii
ACKNOWLEDGMENT.....	iv
TABLE OF CONTENTS .....	v
ABSTRACT.....	x
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
ACRONYMS AND ABBREVIATIONS .....	xiii
<b>CHAPTER ONE: .....</b>	<b>1</b>
<b>INTRODUCTION AND BACKGROUND INFORMATION.....</b>	<b>1</b>
1.0 Introduction.....	1
1.2 Open Access and Digital Preservation.....	3
1.2 History of Digital Preservation .....	5
1.3 The Research Context.....	8
1.6 Aim .....	11
1.7 Objectives of the Study.....	11
1.8 Research Questions.....	12
1.9 Assumptions of the Study .....	12
1.10 Significance of the Study .....	13
1.11 Originality of the research .....	14
1.12 Scope and Limitations.....	17
1.12.1 Scope.....	17
1.12.2 Limitations .....	18
1.13 Definition of Operational Terms.....	19
1.14 Summary.....	22
<b>CHAPTER TWO .....</b>	<b>23</b>
<b>LITERATURE REVIEW AND THEORETICAL FRAMEWORK .....</b>	<b>23</b>
2.0 Introduction.....	23
2.1 Theoretical Framework.....	23

2.1.1 ISO 14721: 2012 Open Archival Information System (OAIS) Reference Model	24
2.1.1.1 Application of ISO 14721: 2012 OAIS Reference Model to the Study	29
2.1.2 ISO 16363:2012 Space Data and Information Transfer Systems -- Audit and Certification of Trustworthy Digital Repository	31
2.1.2.1 Applicability of ISO 16363:2012 to the Study	37
2.2 Need for Digital Preservation	40
2.3 Digital Preservation and Institutional Repositories	44
2.4 Digital Preservation Practices	46
2.4.1 Content in Institutional Repositories	47
2.4.2 File Formats	50
2.4.3 Digital Preservation Policy	53
2.4.4 Digital Preservation Planning	57
2.4.5 Metadata and Digital Preservation	58
2.4.6 Digital Preservation Strategies	60
2.4.6.1 Short-term Strategies	61
2.4.6.2 Medium- to Long-term Strategies	65
2.4.6.3 Investment Strategies	68
2.4.6.4 Alternative strategies	70
2.4.7 Repository platforms and digital preservation	70
2.5 Challenges to Digital Preservation	74
2.6 Digital Preservation Trends	79
2.9 Evaluation of Digital Preservation Practices	81
2.7 Summary	84
<b>CHAPTER THREE</b>	<b>86</b>
<b>RESEARCH METHODOLOGY</b>	<b>86</b>
3.0 Introduction	86
3.1 Research Paradigm	86
3.2 Research Approach	92
3.3 Research Design	96
3.3.1 Case study	97
3.7 Sample Design and Sampling Techniques	101

3.8 Sample Size.....	103
3.9 Data Collection Techniques/Methods.....	104
3.9.1 Face to Face Interviews .....	105
3.9.2 Observation .....	107
3.9.3 Document Analysis.....	108
3.10 Reliability and Validity.....	109
3.11 Data Analysis .....	112
3.12 Ethical issues.....	119
3.13 Summary .....	121
<b>CHAPTER FOUR.....</b>	<b>124</b>
<b>DATA PRESENTATION, ANALYSIS AND INTERPRETATION .....</b>	<b>124</b>
4.0 Introduction.....	124
4.1 Response Rate .....	124
4.2 Characteristics of Materials Accepted into the Repositories .....	125
4.2.1 Types of Digital Materials .....	127
4.2.2 Defined Boundaries for Materials accepted.....	129
4.2.3 Types of Formats Accepted in the Institutional Repositories .....	131
4.2.4 Procedures for Receiving (Ingesting) Materials in the IR .....	135
4.2.5 Metadata.....	138
4.3 Preservation Policy .....	142
4.3.1 Existence of Policy .....	144
4.3.2 Adherence to Policy.....	144
4.3.3 Revision of Policy.....	145
4.3.4 Responsibility for Preservation.....	145
4.3.5 Rights Management .....	146
4.4 Digital Preservation Plan .....	146

4.4.1 Existence of Digital Preservation Plan .....	148
4.4.2 Financial Planning .....	148
4.4.3 Technology Planning .....	149
4.4.4 Staffing and Skills Development .....	149
4.5 Digital Preservation Strategies.....	150
4.6 Challenges of Digital Preservation .....	153
4.7 Proposed Solutions to Challenges in Digital Preservation .....	155
4.10 Summary .....	156
<b>CHAPTER FIVE .....</b>	<b>157</b>
<b>DISCUSSION OF FINDINGS .....</b>	<b>157</b>
5.0 Introduction.....	157
5.1 Characteristics of Materials Accepted in the Repository.....	158
5.1.1 Types of Digital Materials .....	158
5.1.2 Defined boundaries for materials Accepted.....	160
5.1.3 Types of Formats Accepted in the Institutional Repositories .....	161
5.1.4 Procedures for receiving (Ingesting) Materials in the IR.....	164
5.1.5 Metadata.....	166
5.2 Preservation Policy .....	169
5.3 Digital Preservation Plan .....	171
5.4 Digital Preservation Strategies.....	174
5.5 Challenges of Digital Preservation .....	177
5.6 Proposed Solutions to Challenges in Digital Preservation .....	178
5.7 Summary .....	179
<b>CHAPTER SIX .....</b>	<b>181</b>
<b>SUMMARY, CONCLUSION AND RECOMMENDATIONS .....</b>	<b>181</b>
6.0 Introduction.....	181
6.1 Summary of Findings.....	181
6.2 Conclusion .....	189
6.3 Recommendations.....	190
6.4 Proposed Digital Preservation best Practices Framework for Institutional Repositories .....	194



6.5 Recommendations for Further Research.....	201
<b>REFERENCES.....</b>	<b>202</b>
APPENDIX 1: INTERVIEW SCHEDULE.....	232
APPENDIX 2: OBSERVATION AND DOCUMENT ANALYSIS GUIDE.....	234
APPENDIX 3: NACOSTI RESEARCH AUTHORIZATION.....	235
APPENDIX 4: NACOSTI RESEARCH PERMIT.....	236
APPENDIX 5: UON AUTHORITY TO COLLECT DATA.....	237
APPENDIX 6: JKUAT APPROVAL TO CARRY OUT RESEARCH.....	238
APPENDIX 7: SIMILARITY INDEX.....	239

## ABSTRACT

Rapid changes in technology have put a demand on institutional repositories to consider preservation in order to fulfill their goals and objectives. Although research funders, depositors and other stakeholders need evidence that the repository is worthy of trust in relation to long-term preservation, the efficaciousness of these efforts are narrowly documented. This study aimed at establishing the efficacy of digital preservation practices in Institutional Repositories (IRs) of selected public universities in Kenya. Specifically it sought to achieve the following objectives: to determine the types of digital resources in the institutional repositories; appraise the digital preservation plans of the selected IRs; audit the existing digital preservation policies; examine the digital preservation strategies practiced by the IRs in the selected universities; identify digital preservation challenges in the IRs of the selected public universities and to make recommendations and propose best practices for digital preservation in the IRs in Kenya. The study was informed by both ISO 14721:2011: Open Archival Information System (OAIS) model and ISO 16363:2011: Space data and information transfer systems -- Audit and certification of trustworthy digital repository. The research adopted a pragmatic research paradigm and utilized the qualitative research approach incorporating a multi case study research design. Criterion purposive sampling was used to select three universities from which 19 respondents consisting of senior library management, institutional repository staff, systems librarians and ICT staff responsible for ICT issues in the library were sampled using expert sampling. Primary data was collected through interviews, observation and document review with trustworthiness established through triangulation, dependability, confirmability and transferability Data was analyzed using directional content analysis. The findings established that repositories were custodians of both digital and hybrid information resources that ingested into the IR without the support of a needs assessments, inadequate metadata, failure to consider file formats that support long term preservation, lack of comprehensive digital preservation policies, fragmented digital preservation planning, limited poor technology plans and budgets, failure to document preservation actions, limited skills as well as reliance on short term digital preservation strategies. The study concludes that the digital preservation practices could not guarantee long term preservation of the resources in the IRs. The following recommendations are proposed: review of the goals of the IR as this was seen as a major step towards determining the level of digital preservation required, development of: content selection policies, digital preservation policies and plans, development of a business model for digital preservation, investment in stakeholder sensitization and training on digital preservation, formation of cooperatives to address digital preservation challenges and promotion of resource sharing. Finally, framework for best practices in digital preservation informed by the current industry standards is proposed.

**LIST OF TABLES**

Table 1: Research gap and how it is addressed through research questions.....	15
Table 2: Linking the standards to the research questions .....	39
Table 3: Recommended open standard technology neutral file formats.....	51
Table 5: Population type characteristics and size .....	101
Table 6: Sample Population Framework.....	104
Table 7: Themes and subthemes as derived from theoretical framework .....	115
Table 8: Sample directed content analysis.....	118
Table 9: Linkage of research objectives and questions to sources of data .....	123
Table 10: Respondent rate by participant categories .....	125
Table 11: Summary of responses on characteristics of digital materials accepted in the IRS .....	125
Table 12: Distribution of content type across the university .....	127
Table 13: Type of content received in the IRs of the selected universities .....	129
Table 14: File Formats Accepted in the IRs .....	134
Table 15: Guidelines for Ingest as Reported Verbatim .....	136
Table 16: summary of digital preservation policy responses.....	143
Table 17: Summary of responses on adequacy of plan to digital preservation .....	147
Table 18: Summary of responses on adequacy of digital preservation planning .....	150

**LIST OF FIGURES**

Figure 1: OAIS functional model. Adapted from consultative committee space data systems, 2002 .....	29
Figure 2: Trusted Digital Repositories Framework .....	37
Figure 3: Sample Pictorial Collection.....	130
Figure 4: Sample Unidentified File Format .....	132
Figure 5: Sample CDROM storage environment.....	137
Figure 6: Sample metadata collected via dspace .....	140
Figure 7: Sample metadata template.....	141
Figure 8:Sample dead-link message for IR.....	152
Figure 9: Proposed framework for best practices for digital preservation.....	194

**ACRONYMS AND ABBREVIATIONS**

AIP	Archival Information Package
ARK	Archival Resource Key
ASIS&T	Association of Information Science & Technology
AVHRR	Advanced Very High Resolution Radiometer
CCSDS	Consultative Committee for Space Data Systems
CD-ROM	Compact Disk Read Only Memory
CUE	Commission for University Education
DCC	Digital Curation Centre
DOI	Digital Object Identifier
DPC	Digital Preservation Coalition
DPHIL	Doctor of philosophy
EOS	Enabling Open Scholarship
ESRC	Economic and Social Research Council
ICT	Information and Communication Technologies
INASP	International Network for the Availability of Scientific Publications
IR	Institutional Repository
ISO	International Organization for Standardization
JISC	Joint Information Systems Coalition
KLISC	Kenya Libraries and Information Services Consortium
KU	Kenyatta University

LUNC	Land Use and Natural Resources Inventory Project
MPHIL	Master of Philosophy
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NCST	National Council for Science & Technology
NESTOR	Network of Expertise in long-term Storage and Accessibility of digital resources in Germany
NOAA	National Oceanic and Atmospheric Administration
NSF	National Science Foundation
OA	Open Access
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OAIS	Open Archives Information Systems
OCLC	Online Computer Library Center
OpenDOAR	Open Directory of Open Access Repositories
PDI	Preservation Description Information
PLANETS	Preservation and Long-term Access through Networked Services
RAE	Research Assessment Exercise
REF	Research Excellence Framework
RLG	Research Libraries Group
ROARMAP	Registry of Open Access Repositories Mandatory Archiving Policies
SIP	Submission Information Package

TDR	Trusted Digital Repositories
TOVS	Freelance Agency for Television and Film Industry
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNIVAL	Universal Automatic Computer
UoN	University of Nairobi
URL	Uniform resource locator
URN	Uniform Resource Name

## CHAPTER ONE:

### INTRODUCTION AND BACKGROUND INFORMATION

#### 1.0 Introduction

*“Ironically, the faster we progress in terms of technology, the shorter the lifespan of our tools has become. History that was etched into stone walls or tablets millennia ago is still readable in its original format, much more clearly and easily accessible, in fact, than data “saved” a few years ago on a 3½ inch floppy disk”.* Microsoft 2010

The quote above highlights the fragility of digital resources that necessitates its preservation. Digital preservation is crucial to society and it is the responsibility of those who manage digital collections to take steps to ensure that the resources are comprehensible, relevant, and easily accessible to everyone. Digital preservation involves the maintenance and protection of digital objects from threats such as technical malfunctions, media obsolescence, and organizational failures (Rieger, et al, 2022). These efforts are undertaken to ensure that digital objects are authentic, accurate, available, and usable over time. It is especially important to consider information security, privacy, and compliance with policies.

There is enough evidence to suggest that many potentially valuable digital materials have already been lost and organisations have incurred substantial costs to recover these digital contents as observed in the following examples:

The US Census Bureau saved the 1960 Census on Univac paper tapes that could be read only with a UNIVAC type II-A tape drive that became obsolete in mid-seventies to the



extent that when a decision was made to convert them to computer files only two machines were available in the world to read them (Arora, 2009).

According to Chen (2001), NASA/NSF/NOAA rescued valuable 20-year-long TOVS/AVHRR satellite data documenting global warming. This is research that could have had a very high impact not only in terms of money but also as a basis for future research, policy and planning to mitigate the effects of global warming.

The Land Use and Natural Resources Inventory Project (LUNR) sent information to the state of New York Archives in the 80s but had depended on customized software programs to represent and analyze the data and no metadata was provided about the programs required to render them (Waters & Garrett, 1996). Even if the software had been retained, the hardware and operating system needed to run the software were no longer available (Lakshmi & Jindal, 2004). This particular case highlights the important role best practices and specifically technical metadata plays in digital preservation as it helps identify the types of hardware and software and their versions that enable the preservation team to save the data. This was one case where the archive was not able to save the data. Gaur and Tripathi (2012) expounded on the vulnerability of digital information when citing the economist by comparing the survival of the “doomsday” analog book written in 1086 and still found in the British National archives but its digital version in a 12-inch disk is no longer accessible due to hardware incompatibility. From the foregoing, the importance of digital preservation cannot be overemphasized. According to Robertson and Borchert (2014), disasters can take many forms, including fire, flood, tornado, hurricane, earthquake, tsunami, war, computer viruses and human hackers. Instances of deliberate destruction by human beings is evidenced by the Trump

administration's decision to remove peer-reviewed scientific material on climate change from the websites of the Environmental Protection Agency and the departments of the Interior, Energy, Agriculture and State, (Murgu, 2021). Digital repositories are entrusted with valuable digital information that are sometimes unique (Frank,2018). The need to demonstrate that these repositories can be trusted has been a focus of many studies (Altman et al, 2019, Frank, 2018; Maemura et al,2017;Shajitha & Abdul; 2021) yet Abrams (2018) decries a lack of literature on the efficacy of digital preservation efforts. This study attempted to fill this gap by establishing the efficacy of digital preservation practices in selected IRs of public universities. The study establishes that the adoption of best practices in IRs is still a mileage and proposes a framework for best practices that could be used to guide IRs intent on providing long term access to their digital resources.

## **1.2 Open Access and Digital Preservation**

Open access is gaining a momentum because it has been taunted as a solution to high costs of journal subscriptions. Institutional Repositories (IRs) have been described as digital collections capturing and preserving the intellectual output of a single or multi-university community. They provide a critical component in reforming the system of scholarly communication, a component that expands access to research, reasserts control over scholarship by the academy, increases competition and reduces the monopoly power of journals bringing economic relief and heightened relevance to the institutions and libraries that support them (Crow, 2002; Lynch, 2003). Institutional repositories

are not only support mechanisms to import, identify, store, preserve, retrieve, and export an institution's digital assets (Roy, et al, 2012) but also hold the core intellectual assets of a university, characteristics that enable them to be flexible in ways that support the

institution's variety of business processes. Rumsey (2006) has referred to an IR as a means of gathering the entire research output of an institution in one place as well as a record of that institution's research output. Not only is this useful for storage and management purposes, but also for searching. Many universities have adopted IRs as explained by Labrooy (2013) who gives examples of some of the world's most prestigious institutions that have begun implementing Open Access(OA) policies and implementing IRs including Stanford and Harvard universities. Libraries have diversified information resources ranging from digitized collections to the born-digital content. Most IRs have focused on populating and providing access to the resources in their custody with little concern to the long term access to these resources (Adjei, et al, 2019; Robertson & Borchert, 2014). Strong arguments have been put forward against the role of the IR in preservation based on the fact that earlier repositories like e-prints held surrogates for resources with an analog equivalent, (Pinfield & James 2003). In response to this, Ball (2010) put forward his argument that IRs hold more than information surrogates and should be involved in preservation. The Directory of Open Access Repositories (Open-DOAR) as of September 2022 boasts of five thousand nine hundred and thirty-two (5932) institutional repositories forty-six (46) of these in Kenya (Open-DOAR, 2022). Of importance is the fact that most of these research outputs held by the repositories have long term value (Gbaje & Mohammed, 2017) and are in an ever greater variety of formats, many of which are relatively short-lived (Rauch, et al, 2007). Newer versions of the same software may be unable to render files produced with older versions of that software (Vij, 2017). Soft and hardware environments change constantly and after only a few years, older files often can no longer be rendered with up-to-date systems.

This dictates for a long term plan to ensure they are available and accessible taking into account the very dynamic changes in electronic information storage. Managers of IRs naturally have a responsibility for the longevity of the materials they are charged with managing for their institutions and researchers (Hockx-Yu, 2006). Durant, (2010) defines digital preservation as “the whole of the principles, policies, rules and strategies aimed at prolonging the existence of a digital object by maintaining it in a condition suitable for use, either in its original format or in a more persistent format, while protecting the object’s identity and integrity, that is, its authenticity”(p.1). Scholars like Kirchhoff, (2008) support this holistic definition by describing digital preservation as “the series of management policies and activities necessary to ensure the enduring usability, authenticity, discoverability and accessibility of content over the very long term” (p. 287). The two definitions agree that digital preservation is holistic and takes into consideration management (rules and policies among others) as well as actions/activities necessary to ensure the resources have integrity and are available and usable in future. Preservation and access go hand in hand. It is impossible to promote access without preservation. IRs have to take digital preservation seriously to ensure that their responsibilities to both institutions and users are fulfilled.

## **1.2 History of Digital Preservation**

Prior to the 1990s, digital preservation initiatives were isolated and undertaken by individual organisations and for small record groups or specific records (Baucom, 2019). Wide spread access to the internet and more use of personal computers led scholars like Zweig (1993) as cited in Baucom (2019) to raise concerns regarding the preservation strategies being applied then on the future of research resources. Zweig, acknowledged

that most of these resources remained unprinted and thus remained available only within the institutions, a situation that in future could result in unavailability of the information resources. As a Jewish history researcher, he had used some internal documents that had not been considered for printing and therefore he recommended the need for a comprehensive policy supported by the government that would ensure that these resources were not lost to the future researchers. Digital preservation gained momentum after the 1996 report “Preserving Digital information” commissioned by the Commission on Preservation and Access and the Research Libraries Group (RLG) which despite acknowledging the benefits of digital information such as easy access and multiple sharing, it also recognized the challenges brought about by technological obsolescence, media fragility and machine dependency that posed a threat to long term access if not well taken care of (The Commission on Preservation and Access and the Research Libraries Group, 1996).

Pinfield and James, (2003) raised the question on the necessity of IRs to undertake digital preservation responsibilities. Lynch (2003) felt that a key part of an IR’s service is to manage technological change and the migration of digital content from one set of technologies to the next in order to cope with technological obsolescence. Lynch’s idea is supported by Hitchcock et al (2005) who note that IR managers have a responsibility to ensure the longevity of the materials they are charged with managing for their institutions and researchers. Other researchers argued that an institutional repository’s main responsibility is to broaden distribution and access (Ball, 2010). Similarly, Owens (2018) underscore the need to look at repositories holistically describing a repository as “the sum of financial resources, hardware, staff time, and ongoing implementation of policies and

planning to ensure long-term access to content” (p,7). Shajitha and Abdul (2021) adds that the responsibilities of IRs do not end with their establishment but rather begins there, and proposes an evaluation of its major activities to identify non conformities. Literature however shows a different scenario. For example, Hurley and Shearer (2019) in their report on Digital Preservation Capacity and Needs at Canadian Memory Institutions noted that although 94% of the institutions had digitized their resources, majority of them (85%) focused on providing access. Only 17% had digital preservation policies with only 1% having staff dedicated to digital preservation. Digital preservation is a product of people, institutions, infrastructure, and procedures (Wilczek & Glick, 2006).

Trusted digital repositories perform preservation functions. This notion was first introduced in a seminal report, *Preserving Digital Information* by Waters and Garrett in 1996. The authors emphasized the role of trust in managing the identity, integrity, and quality of digital information in archival systems and recommended developing a process of certification. The concept of a trusted digital repository was fully articulated in another foundational report, *Trusted Digital Repositories: Attributes and Responsibilities*, prepared by a RLG/OCLC working group (Beagrie et al., 2002). A trusted digital repository is defined as one “whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future” (Beagrie et al., 2002, p. 5). In order to gain recognition as “trusted,” a repository has to have certain attributes that ensure the reliability and authenticity of stored information. Donaldson (2020) articulates the need to know whether digital repositories can be trusted with research findings and data by exhibiting characteristics such as: defining the scope of the collections, defining the needs of stakeholders, adopting a lifecycle model for

managing and preserving digital resources, demonstrating institutional viability and operational responsibility by adopting and following procedures and policies among others. Infact, Shajitha and Abdul (2021) aptly conclude that adoption of digital preservation policies and plans lays down the foundation for successful institutional repositories.

The need to create digital repositories that could be trusted with research findings resulted in the development of two foundational international standards on which the digital preservation community still relies on today: The Open Archival Information System (OAIS) model and the Trustworthy Repositories Audit and Certification (TRAC) checklist, which later became ISO 16363, Audit and Certification of Trustworthy Digital Repositories (Baucom, 2019). Baucom (2019) advises digital repositories to use the OAIS model and ISO 16363 for self-assessment when establishing digital archives.

### **1.3 The Research Context**

The International Network for the Availability of Scientific Publications (INASP) funded two members of the Kenya Library and Information Services Consortium (KLISC) to attend an OA workshop and a one-week attachment at the University of Pretoria in 2009. In 2010, Electronic Information for Libraries (EIFL) began advocating for open access in Kenya by sponsoring the first open access workshop (EILF, 2018). In 2016, EILF and KLISC collaborated on another project and funding from SPIDER (the Swedish Programme for ICT in Developing Regions DSV, Department of Computer and System Sciences, Stockholm University) to develop institutional OA policies requiring the deposit of all research output, such as journal articles, theses, and dissertations, in institutional OA repositories (EILF, 2018).

The Commission on University Education (CUE) was not left behind. According to their standards and guidelines, a university cannot be accredited unless it provides a library and information services that meet its defined criteria. Among these criteria is the requirement that university libraries establish and maintain institutional repositories as well as provide for the digitization of information resources (CUE, 2013).

Jain (2012) describe two approaches to implementing open access: the gold approach and the green approach. The gold approach entails publication in open access journals, whereas the green approach entails self-archiving in digital repositories. Both approaches have been adopted in Kenya, but green open access has seen the implementation of a number of repositories in various institutions. The failure by the government to provide enough funding has been credited with the rapid adoption of open access. The traditional model of scholarly communication to disseminate research has become too expensive for anyone library to subscribe to. This coupled with increase of information materials in regard to content and formats has seen the wide adoption of open access by libraries (Wakeling, et al, 2019). Based on the important role open access plays in Kenya, the need for effective digital preservation cannot be underestimated. Mosei (2016) notes that researchers and institutions in Kenya recognize the need for long term accessibility of their research output but they lacked trust in the ability of digital archives to facilitate it. She decried the need for more concerted efforts especially on funding and creating awareness aimed at guaranteeing effective preservation of data. From the aforementioned, it is clear that IRs in Kenya manage valuable content for their designated communities and have a commitment to preserve the digital content to perpetuity. By conducting this study, the researcher aimed to gather valuable data and insights that



informed recommendations and strategies for improving digital preservation practices in Kenyan universities. This, in turn, can help ensure the long-term accessibility and usability of research output, addressing the concerns highlighted by Moseti (2016).

#### **1.4 Statement of the Problem**

Studies have revealed that the ingestion of digital files into the IR, is not a guarantee that they will be accessed for use over long-term (Hurley & Shearer, 2019) neither is it sufficient to generate the trust of users (Anderson, 2015). Correspondingly, Anderson (2015) accentuates the need for archives to exhibit active preservation for enduring value of the resources in their custody. Digital preservation is arguably seen to be much more than merely avoiding loss with Frank (2018) proposing that digital repositories should acquire sustainable infrastructure that includes institutional frameworks, fiscal sustainability, and sound procedures that ensure viability and accessibility of the digital resources in the longterm. Arguably therefore, frequent assessment of institutional repositories and organization's abilities to achieve their digital preservation goals is core to robust digital preservation (Frank, 2018; Maemura et al, 2017;Shajitha & Abdul, 2021). Moreover, Tieman (2015) point out that research funders, depositors and other stakeholders need evidence that the repository is worthy of trust.

Yet, as a 2018 survey of the National Digital Stewardship Alliance (NDSA) member institutions revealed, the identification and evaluation of effective digital preservation practices is a most important challenge within the community of digital stewards (Altman et al, 2019). Abrams, (2018) decries the lack of literature on the success of digital preservation practices and opines that without knowing the efficacy of digital preservation efforts, it is impossible to plan for it. As a counter measure, Donaldson

(2020) proposed the use of a systematic and independent audit to determining the details of the digital preservation process and identify potential weak points in order to make improvements.

Institutional repositories in Kenya have adopted several strategies to ensure that the digital resources in their custody are availed to future generations (Erima, Et al, 2016; Moseti, 2016). The dearth of literature on the effectiveness of these strategies is not only a critical knowledge gap, but also it denies interested stakeholders such as the librarians who are the chief custodians, the faculty members who are the primary contributors, and the parent organisations and any other funders, evidence-based empirical feedback of the efficacy of the preservation measures. It is against this background that this research seeks to answer the question, “are the digital preservation practices in IRs in Kenya effective enough to support long term preservation?”.

### **1.6 Aim**

This study aimed at evaluating digital preservation practices in Institutional Repositories (IRs) of selected public universities in Kenya with a view to proposing recommendations as well as a best practices framework for effective digital preservation in IRs.

### **1.7 Objectives of the Study**

Specifically, the research seeks to:

1. Determine the characteristics of digital resources in the institutional repositories.
2. Audit the existing digital preservation policies
3. Appraise the digital preservation plans of the selected IRs

4. Examine the digital preservation strategies practiced by the IRs in the selected universities
5. Identify digital preservation challenges in the selected public universities.
6. Make recommendations and propose best practices for digital preservation in the IRs of the selected universities.

### **1.8 Research Questions**

1. What are the characteristics digital objects accepted for storage in IRs in the public universities under study?
2. How suitable are the digital preservation policies in supporting digital preservation?
3. How adequate are the digital preservation plans of the selected IRs?
4. What preservation strategies are used by the IRs and how adequately do they address long term preservation?
5. What challenges do the IRs experience in preserving research outputs?
6. What can be done to improve the effectiveness of digital preservation practices in the selected universities?

### **1.9 Assumptions of the Study**

This study was carried out with the following assumptions, that:

- i. Transparent and well-documented digital preservation practices contribute to the epistemic reliability and trustworthiness of digital materials in repositories.
- ii. Digital materials within institutional repositories are of enduring value, and their preservation is a reflection of the IR's commitment to the user community.

- iii. Universities are actively preserving their own research output, including academic publications, theses, and dissertations, in their institutional repositories.

### **1.10 Significance of the Study**

Institutional repositories play a critical role in scholarship especially in developing countries where funding is reducing as time goes. Research on the management of the resources under their care assists the IR managers to achieve this role. This study fulfils this role by:

#### **Theoretical Significance**

This research contributes to the existing body of knowledge on the preservation of research outputs in institutional repositories by providing empirical evidence on practices, challenges and concerns that need to be addressed to ensure that long term access is possible. Researchers could also utilize it to build on digital preservation research.

#### **Policy Implications**

The findings will create awareness among university and library management on the strategic role of digital preservation to the IRs' overall goals so that it is incorporated into their strategic plans. It also hoped that the findings and recommendations will inform digital preservation policies and procedures on sustainable digital preservation as adoption of a life cycle concept in the management of digital resources.

#### **Practical Significance**

The outcome of this thesis is a framework of best practices that IRs can adopt to establish robust digital preservation programs that will adequately address the trust issues of their users.

### **1.11 Originality of the research**

A number of studies have been carried out in the aspect of digital preservation in Africa. Studies by Moseki (2016) and Erima, et al (2016) in Kenya have focused on digital preservation practices within universities and identified the strategies utilized by the universities to preserve the information. Barrueco and Termens (2022), decry a lack of evidence in the literature on how IRs (institutional repositories) are following through on their commitment to provide long-term access to digital content, despite their stated intent to do so. A number of studies have been done with a focus on determining the effectiveness of digital preservation practices. Masenya and Ngulube (2019) identified the need for the adoption of best practices that will support long-term digital preservation of resources held by academic libraries in Nigeria while Umana's (2020) research within IRs in universities in Namibia investigated the long-term digital preservation activities with the findings indicating serious IRs gaps between best practices and what was being done in terms of policy, skills development and management support. Adjei, et al, (2019) were very categorical that the practices utilized by digital repositories in Ghana were very inadequate to the extent that if nothing was done the IRs risked loss of information in their custody. Research focusing on establishing efficaciousness of digital preservation practices in Kenya are limited. Evaluation is context specific thus the need to establish whether digital preservation practices within the institutional repositories support long-term preservation. This research aims to fill this gap in Kenya. Table 1 provides a summary of the research gap. Table 1 summarizes the gaps identified from literature.

**Table 1: Research Gap and how it ss Addressed through Research Questions**

Research gap	How the gap is addressed	Research question
Smith (2008) recommended a needs assessment to meet the needs of the designated community. Barrueco & Termenes (2022)There is a focus on situation in libraries and less on preservation metadata. Altman et al (2019) organisations have lost much content due to organizational failure; proposes categorization of content.	Determine the characteristics of digital resources in the institutional repositories.	What types of digital objects are accepted for storage in IRs in the public universities under study?
Policy development is a vital digital preservation strategy (Faundeen, 2017;Ismail & Affandy, 2018). Dell & Shultz (2014) presence of a digital preservation policy is an indication that an organisation commitment to practice digital preservation. ISO 14721:2012, a digital archive should have a policy that addresses metadata, content selection, strategies, data security and defined file format policies.	Audit the existing digital preservation policies	Are the digital preservation policies comprehensive enough to support digital preservation?
ISO 14721:2012 proposes the audit of digital preservation practices including digital preservation planning. International Records Management Trust (2016) established a general lack of awareness on digital preservation planning that resulted in its exclusion from the overall organizational plans	Appraise the digital preservation plans of the selected IRs	How adequate are the digital preservation plans of the selected IRs?
Rieger (2022) institutions are focusing on preserving digitized content but lack adequate strategies for born digital content. ISO 16363:2012 frequent evaluation of	Examine the digital preservation	To what extent do the preservation strategies being used by the IRs adequate

<p>digital preservation strategies should be to ensure that they support long-term preservation;</p>	<p>strategies practiced by the IRs in the selected universities</p>	<p>long term preservation?</p>
<p>Johnson (2020) decries there many types of file formats, storage media, hard ware and software that no longer exists; Langley (2019) warns of the short time available to save digital resource due to obsolescence;</p>	<p>Identify digital preservation challenges in the selected public universities.</p>	<p>What challenges do the IRs experience in regard to preserving research outputs?</p>
<p>Umana (2020); Adjei, Mensah and Amoafu (2019) Identified serious gaps in adoption of best practices. Masenya and Ngulube (2019) opined the need for the adoption of best practices that will support long- term digital preservation of resources held by academic libraries in Nigeria</p>	<p>Make recommendations and propose best practices for digital preservation in the IRs of the selected universities.</p>	<p>What can be done to improve the effectiveness of digital preservation practices in the selected universities?</p>

## **1.12 Scope and Limitations**

### **1.12.1 Scope**

The research was undertaken in three public universities derived from a sample of public universities that had made their repositories open as reflected by their registration in open DOAR. The research targeted public universities on the argument that being publicly funded, they are obliged to make their research findings available for free to the public and to preserve the same for posterity. Three universities formed the cases of the study were chosen based on the extent of development of their institutional repository infrastructure defined by the duration the repository has been registered with Open ROAR (e.g., UoN in 2013, KU in 2012, JKUAT in 2013), the volume of content available on their repository websites at that time (comprising over 3500 items), and their registration with ROARMAP. The criterion for evaluating the registration time was based on the assumption that the universities had established robust institutional repositories with substantial content, possibly requiring active digital preservation efforts. Additionally, registration with Open ROAR signified the presence of policies, which was one of the aspects being investigated. The research focused on examining the digital preservation practices within the libraries of these universities since they host the IRs. Nonetheless, the ICT departments that advise the library on hardware and software issues including ICT policies were included.

The research is limited to the preservation of only information resources such as documents, videos and images and will not include the preservation of software and databases although it acknowledges that there also exists a research gap in this area too.



The study was also limited to examining the efficacy of digital preservation practices from a managerial and program perspective. The successful use of the preserved digital resources by the designed community is also important in the establishment of the success of a digital preservation program. However, this was not a focus of this study.

The respondents were drawn from senior library management because of their strategic position in the drafting and implementation of policies and procedures, the repository personnel, and the library ICT link persons.

### **1.12.2 Limitations**

Although the study focused on public universities selected on the assumption that their institution repository infrastructure was well developed the findings may not be generalizable to private universities, other countries, or institutions with significantly different resources and contexts.

The study's focus on selected public universities in Kenya narrows down the scope significantly. While this focus allows for in-depth analysis within the chosen context, it does not encompass the broader landscape of digital preservation practices in other types of educational institutions, private universities, research institutions, or non-academic organizations in Kenya. Therefore, the findings provision of a comprehensive view of digital preservation practices across diverse sectors and may not be applicable to institutions beyond the specific universities studied. This limitation restricts the study's ability to provide a holistic understanding of the digital preservation landscape in Kenya.

### 1.13 Definition of Operational Terms

**Bit Stream:** Bit stream refers to a contiguous or non-contiguous data within a file that has meaningful common properties for preservation purposes.

**Content Information:** Content information refers to a set of information originally targeted for preservation and is made up of the content data and its metadata.

**Corrigenda:** changes to the article an author may wish to publish at any time after acceptance

**Digital Preservation Plan:** A series of preservation actions to be taken by a responsible institution due to an identified risk for a given set of digital objects or records called collection.

**Digital Preservation Policy;** A plan of action for the safekeeping of digital objects that addresses the questions of what needs to be preserved, why, for what purpose, and for how long.

**Digital Preservation:** A combination of policies and work flows that support the active management of digital objects to ensure their continued authenticity and meaningful access through time and changes in technology

**Emulation:** A means of overcoming technological obsolescence of hardware and software by developing techniques for imitating obsolete systems on future generations of computers.

**Migration:** A means of overcoming technological obsolescence by transferring digital resources from one hardware/software generation to the next with the purpose of

preserving the intellectual content of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology.

**Errata:** correction of errors introduced to the article by the publisher often done during proof reading

**Ingest:** To accept one or many submission information packages (SIPs) into an Archive

**Institutional Repository:** Digital collections that capture and preserve the intellectual output of a single or multi-university community.

**Long Term:** Long-term in digital preservation has been taken to mean as long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long term may extend indefinitely.

**Long-term digital preservation (LTDP):** Series of managed activities required to ensure continued access to digital materials for as long as necessary.

**Metadata:** Metadata is structured information that describes, explains, locates or otherwise makes it easier to retrieve, use, manage and preserve a digital information resource.

**Open Access:** Open access is a new model of scholarly communication through which the author(s) and right holder(s) of scholarly work grant(s) to all users a free, irrevocable, worldwide right of access to, and a license to copy, use, distribute, transmit, and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper

**Open Archival Information System:** A type of archive consisting of an organization of people and systems that has accepted the responsibility to preserve information for one or more designated communities.

**Reformatting:** Reformatting means copying information content from one storage medium to a different storage medium (media reformatting) or converting from one file format to a different file format (file re-formatting).

**Refreshing:** This means copying information content within the same storage media.

**Rendered Digital Objects:** Digital objects which are processed by some software to produce a rendering which is presented to a human user who can then interpret whathe/she sees/hears/feels/tastes and include documents, pictures, videos and sounds. Include the terms that I indicated in the document.

**Sustainability:** A set of business, social, technological, and policy mechanisms that encourage the gathering of important information assets into digital preservation systems, and support the indefinite persistence of digital preservation systems, enable access to and use of the information assets into the long-term future

**Trusted Digital Repository:** A digital repository whose mission is to provide long-term access to managed digital resources; accept responsibility for the long-term maintenance of digital resources; designs its system(s) in accordance with commonly accepted conventions and standards to ensure the ongoing management, access, and security of materials deposited within it; establishes methodologies for system evaluation that meet community expectations of trustworthiness; can be depended upon to carry out its long-

term responsibilities to depositors and users openly and explicitly; and whose policies, practices, and performance can be audited and measured

### **1.14 Summary**

This introductory chapter provided background information to the research and discussed the initial stimulus for the study. The chapter establishes that Institutional repositories play a critical role in the scholarly communication process. It also identifies two major aims of the IRs: To disseminate research findings and to guarantee their long-term access. However, repositories focus on the first aim and rarely on the second.

The need for IRs to address digital preservation is brought out by examples of digital content loss due to limited digital preservation actions. It also brings out the relationship between digital preservation and access. The critical role of digital preservation plans and policies is established.

It is established that open access publishing has been adopted in Kenya with scholarly content being deposited in the IRS. The chapter has also established that these IRs are practicing digital preservation.

This study aimed at evaluating digital preservation practices in Institutional Repositories (IRs) of selected public universities in Kenya with the aim of proposing recommendations as well as best practices for effective digital preservation in IRs. Existing studies have established that IRs have some digital preservation actions in place but the extent to which these practices support long term access reveals limited to literature. This is the gap this study is focusing to fill.

## CHAPTER TWO

### LITERATURE REVIEW AND THEORETICAL FRAMEWORK

#### 2.0 Introduction

This chapter reviews literature of previous research on institutional repositories and digital preservation. It introduces the framework for the case study that comprises the main focus of the research described in this thesis. Literature has been reviewed along the lines of purpose of digital preservation, digital preservation in institutional repository, digital preservation practices (content, digital preservation planning, digital preservation policies, digital preservation strategies), current digital preservation initiatives and evaluation of digital preservation practices. The chapter begins with review of the theoretical literature before proceeding to the empirical literature.

#### 2.1 Theoretical Framework

The purpose of this study was to assess digital preservation practices in institutional repositories of selected public universities in Kenya. The success of any digital preservation program, according to Abrams (2018), can be measured on two levels: the trustworthiness of managerial systems and programs, and the successful use of preserved resources. Three international standards have been identified as tools that can be used in the evaluation of digital repositories. These are: ISO 14721:2012: Space data and information transfer systems -- Open archival information system (OAIS), ISO 16363: 2011: Space data and information transfer systems -- Audit and certification of trustworthy digital repositories and ISO 16919: 2012: Space data and

information transfer systems -- Requirements for bodies providing audit and certification of trustworthy digital repositories (Downs, 2019).

This study utilized ISO 14721:2012, Open Archival Information System (OAIS) model and ISO 16363: 2011, Space data and information transfer systems -- Audit and certification of trustworthy digital repositories. ISO 16919 is used by the organisations that audit and certify trusted digital repositories and therefore not applicable to the study. OAIS reference model was developed to assist organizations in answering questions in regard to the creation, deposit, archival storage and use of the digital objects that the organization preserves (Allinson, 2006). OAIS is a reference model that guides the development of sustainable digital preservation programs by providing a common vocabulary, an information model, and a high-level digital preservation architecture (Zierau, 2017). It is a standard that has been adopted by many organisations today to inform these organisations on how to establish and manage open archives (Lee, 2005). Some of the organisations that have been able to adopt it include the US National archives (NARA), the Library of Congress, the British Library, Digital Curation Centre in UK, Jstor for journal storage and many more, (Lee, 2005).

### **2.1.1 ISO 14721: 2012 Open Archival Information System (OAIS) Reference Model**

The OAIS provides a definition of a reference model as “A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment” (ISO 14721, 2012). A reference model is based on a small number of

unifying concepts and may be used as a basis for education and explaining standards to a non-specialist (ISO 14721, 2012).

Many arguments have been put forward explaining how a digital archive can demonstrate fitness to facilitate long-term preservation of digital resources. Notable amongst them suggest that a digital archive should be able to: fulfil mandatory requirements; adopt the information model; adopt the terms and concepts as defined by the OAIS reference model for its documents and standards as well as be able to fulfil all the functions described in the model because the mandatory responsibilities cannot be fulfilled if the functions are not performed as described (Allinson 2006; Beedham et. al 2005;Lavoie 2004).OAIS requires a digital repository to negotiate and accept appropriate information from information producers in order to ensure that it has control of what is acquired for successful digital preservation. A number of studies have recognized the importance of this responsibility. Baucom (2019) proposes that OAIS archive should define the type of data to be collected as well as set procedures for negotiating for the information to obtain control of information in order to obtain permissions for preservation. Beedham, et al (2005) are of the opinion that repositories should have supporting policies and procedures that govern the selection criteria for the content that will be accepted in the repository. Hedstrom (1998) and Baucom, (2019) hold that user needs and preservation are inseparable and that information organisations cannot accomplish their preservation missions if their users' needs are not considered. Some authors have also suggested that selection of digital resources for preservation should be done at the early stages of their lives (Conway, 2000; Lee & Tibbo, 2007). The OAIS reference model further requires that



a digital archive should obtain sufficient control of the information to ensure long-term preservation because intellectual property rights have been identified as a major challenge for digital preservation.

According to the OAIS framework, a compliant repository should determine the scope of the designated community. ISO 14721 (2012) defines designated community as “an identified group of potential consumers who should be able to understand a particular set of information”. Consequently, Bountouri, et al,(2018) have argued that defining the designated community for each set of information that has to be preserved is significant, since it influences the selection of the content to be included in the repository and the creation of metadata that will represent the information in order for the designated community to be able to interpret the data. Identification of the designated communities is not without challenges. Talboom and Underdown (2019) report that most institutions faced challenges when identifying designated communities especially those serving a variety of users as well as those faced with born digital resources that could be represented in a myriad number of ways. The responsibility of preserving the information so that it is understandable by the designated community places demands that digital repositories utilize documented policies and procedures to ensure the integrity of the digital resources in their custody (Bettivia, 2016). Under the model, the repository is also required to put mechanisms in place to support digital rights management for copyrighted materials in their custody (Kastellec, 2012).

The responsibilities articulated by the OAIS model encompass activities that repositories undertake on a day today basis. Solicitation for deposits, promoting the

use of the repository, establishing procedures for ingest and engaging management in policy development and projects for financial sustainability are activities that a digital archive would not be functional without (Beedham et al, 2005). Baucom (2019) recommends that digital repositories should have transparent policies and procedures to guarantee the long-term preservation of and access to the digital objects in their custody.

The model goes beyond the responsibilities and defines a set of functions for packaging the digital content as it moves from creator to end user and as it is preserved for long term access. The ingest function receives information from producers and prepares it for storage and management within the archive in form of Submission Information Package (SIPs). Quality control actions such as checking for completeness, technical control of files including formats, and presence of malware as well as format conversion is done during ingest (Korb & Strodl, 2010).

The archival storage function handles the storage, maintenance and retrieval of Archival Information Packages(AIPs) held by the archive. Fleischhauer, et al (2000) describes AIPs as the digital equivalents of archival items such as a books, record albums, or motion pictures consisting of the digital content and metadata that describes the structure, content, and meaning of the data files that has been (encapsulated) either logically or physically as an entity. Migrating digital objects to new media, error checking, implementing disaster recovery strategies, and providing copies of requested AIPs to the access function are some of the activities undertaken within this function, (ISO: 14721, 2012;Tomášek, 2018).

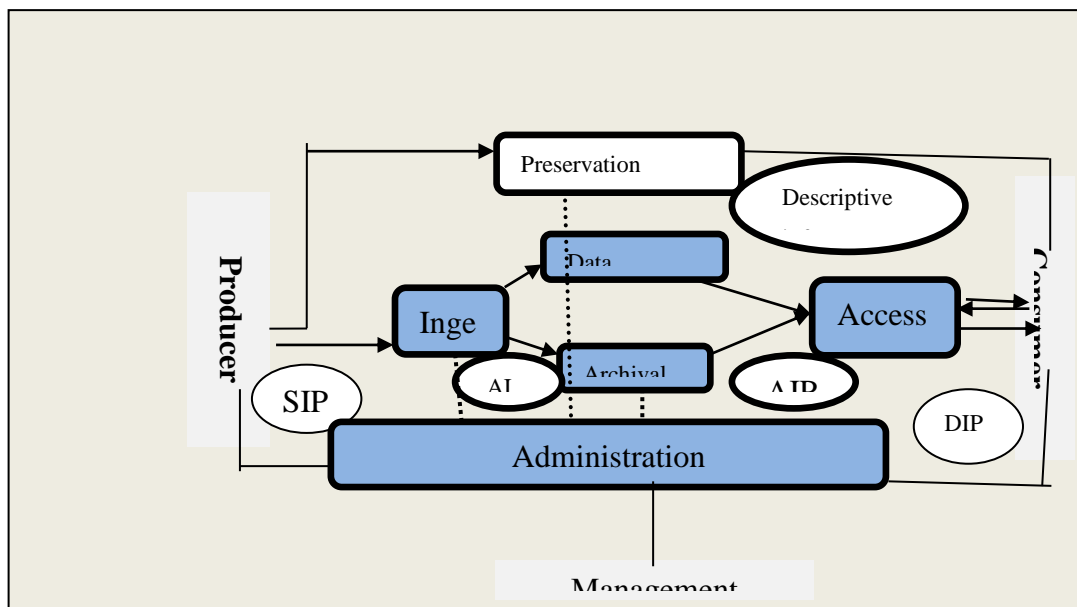
Preservation planning function involves constant monitoring of the digital preservation landscape in order to effect changes to the preservation program that will comply with international standards and best practices (Baucom, 2019). According to Andriamahady, (2021) the plan recommends information updates, migration recommendations, periodic risk analysis reports, as well as technology innovations. Preservation planning ensures that all tasks involved in keeping the digital material accessible and understandable in the long term even in the face of technology obsolescence are supported (Rahmanto & Riasetiawan, 2018).

The data management function coordinates descriptive information pertaining to the archive's AIPs, in addition to system information used in support of the archive's operation. In a nut shell the function entails activities such as database updates, performing queries on database, database administration as well as generation of reports, (Lee, 2010). Data management requires a good inventory that captures and identifies descriptive (provenance, title, date), technical (file name, and file format as well as file size) organizational (access rights, copyrights) metadata, (Goldner, 2017).

The administration function manages the day-to- day operations of the archive. This includes negotiating submission agreements with information producers and performing system engineering, access control and customer services as well as the performance regular audits of to establish compliance with the submission agreement in addition to the development policies and standards (Lee, 2010). This function also serves as an interface between the archive and two components of the OAIS environment: management and the designated community (see Figure 1). It has been taunted as the most complex function as it incorporates both technical and human

processes such as audit, policy making, strategy as well as customer service (Allinson, 2006).

The access function helps consumers to identify and obtain descriptions of relevant information in the archive, and delivers information from the archive to consumers. The function controls the user's ability to search and use information resources by providing the interface through which the end user interacts with the system, (Lee, 2010). The access function calls for a user interface that can be interrogated by users to find, locate and use resources from the archive (Breeding, 2002). Figure 1 below outlines the functional components of an OAIS archive.



*Figure 1: OAIS Functional Model. Adapted from Consultative Committee Space Data Systems, 2002*

### 2.1.1.1 Application of ISO 14721: 2012 OAIS Reference Model to the Study

Taking into consideration the objectives of this study, the OAIS model was found suitable as a tool for establishing the efficacy of digital preservation practices in

institutional repositories under study since by defining the minimum requirements and the roles of different stakeholders it spoke to the first objective of this study that examined the characteristics of the digital resources within the institutional repositories. This objective was also covered by the framework's information model that defines the kind of metadata a digital archive intent on long-term digital preservation should keep. Corresponding to the type of content in the IRs, OAIS requires institutional repositories to negotiate for relevant content by defining the type of data (both in content and file format) accepted in the repository as well as providing procedures for those responsible for the uploading the digital objects.

Additionally, the model emphasizes the development of policies and plans that support preservation and long-term access under its administration function which resonates well with the second and third objective of this study which sought to audit the existing digital preservation policies and establish the existence of digital preservation planning in the selected public universities in Kenya respectively.

Lastly, the model supports the fourth objective which aimed at auditing the digital preservation strategies practiced by the IRs in the selected universities through its archival function that recommends that digital repositories should adopt digital preservation strategies that ensures long-term access to information resources by the designated community.

The choice of the model was also anchored on Thibodeau's (2007) assertion that it presented itself as a benchmark for the evaluation of digital repositories. ISO 14721:2012 therefore provided evaluation guidelines for the study.

The OAIS model however, was not able to wholly support the achievement of the goals of this research since it only articulates the requirements of a digital archive but lacks the metrics to evaluate the effectiveness of the day to day practices. Lin et al (2020) notes that conforming to the OAIS reference model does not guarantee trustworthiness and recommends additional elements such as appropriate governance, resources, and security. In addition, OAIS is a reference model and does not provide detailed implementation guidelines making it prone to different interpretations and implementations necessitating audit and certification (Lin, et al, 2020). To overcome these challenges and achieve all the objectives of the study, ISO 16363: Space data and information transfer systems -- Audit and certification of trustworthy digital was used to address these gaps.

### **2.1.2 ISO 16363:2012 Space Data and Information Transfer Systems -- Audit and Certification of Trustworthy Digital Repository**

The idea of certifying trusted repositories is traced back to a 1994 report by the Commission on Preservation and Access (CPA) and the Research Libraries Group (RLG) (Baucom, 2019).

The task force argued that copyright law limited the certification of digital preservation rendering it difficult for repositories to entrusted the rights for storing and providing long-term access to digital objects. In their view, certification would repositories would expand the repositories legal rights to handle digital objects including those that are not originally owned by the repository. It was this accentuate on certification that gave rise to the concept of 'Trusted Digital Repositories'(TRD).

In 2002, *The Trusted Digital Repositories: Attributes and Responsibilities (TDR)* was jointly published by Research Libraries Group (RLG) and the Online Computer Library Center (OCLC) with a view to articulating a framework of attributes and responsibilities for trusted, reliable, sustainable digital repositories which were required for an archive to provide permanent or indefinite long-term preservation of digital information. Research Libraries Group (RLG), (2002) describes a trusted digital repository as one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. Anderson (2015) declares that providing access to digital resources by digital resources was inadequate to gain the trust of users and instead urged the repositories to demonstrate that they were actively preserving the resources entrusted to them by the users. Lin et al (2020) declared that repositories had to earn the trust of their users and in addition demonstrate that they can be relied upon and at the same time are capable of appropriately managing the data they hold. According to Dobratz, et al(2007) the trustworthiness of a digital repository entails being able to operate according to its objectives and specifications. This means that the evaluation of a digital repository's effectiveness is measured against its goal and mandate as stated in its mission statement. Preservation of digital objects is a process and its evaluation should be done on all the steps involved in the establishment of a digital repository (Dobratz, et al 2007). The TRD framework was developed to help digital repositories and bodies involved in certification to evaluate whether a repository qualifies to be a trusted repository. In 2003, the Research Library Group created a joint task force with the National Archives and Records Administration (NARA) to develop specific

criteria that would “facilitate the certification of digital repositories,” as defined by the report on Trusted Digital Repositories, (Baucom, 2019). This taskforce outlined seven attributes that a repository had to meet to be trustworthy. These attributes are: Compliance with the OAIS reference model, administrative responsibility, organizational viability, financial sustainability, technological and procedural suitability, system security, and procedural accountability. The TRD framework became ISO 16363/TDR 2011.

The administrative responsibility attribute requires that a trusted digital repository provides evidence that it is committed to adopt industry standards and best practices in regard to those that influence its operations especially those that directly influence its viability and sustainability. These include compliance to OAIS as well as standards relating to the physical environment, backup and recovery procedures, and security systems. This responsibility extends to meeting stakeholder expectations and frequently carrying out quality assessments and communicating the same to the customers (RLG, 2002). In addition to these responsibilities, the trusted digital repository is required to involve external auditors to help validate their processes and procedures on a regular basis. Written agreements with depositors and ongoing risk management and contingency planning should be part of the organization’s annual strategic planning.

Consequently, organizational viability relates to the organizational environment surrounding the digital archive. Faundeen (2017) emphasizes the importance of a good organizational environment for the viability of a digital repository. According to the ISO 16363, organizations choosing to become trusted digital repositories should



establish themselves in ways that demonstrate their viability. Aspects such as mission and goals (Zuccala, et al, 2008), policies and procedures that are continuously reviewed (Bantin, 2016; Faundeen, 2017), specialized resources, and adoption of standards (Dollar & Ashely, 2020) as well as digital preservation strategic plans (Bantin, 2016;Lampert & Vaughan 2018) demonstrate an organization's viability.

The ISO 16363 goes on to add that a trusted digital repository should be able to prove its financial sustainability over time and adhere to good business practices including a business plan and adequate budget and reserves. The National Information Standards Organization(NISO) (2010) identified financial commitment as one of the major challenges digital repositories are unable to address especially those associated with the government funding due to the fact that activities involved may not fit in an annual budget. ARMA (2015) recognize the importance of this aspect and postulate that long term digital preservation is expensive not only due to infrastructure required but also funds for staffing and technology watch. NISO (2010) goes on to argue that although this is a challenge, organisations should show commitment by use of mandates, membership subscriptions for collaborations, fundraising and endowment funds. The Cornell University Library (2010) recognized the lack of funding models to support financial sustainability and recommended more work to identify workable solutions especially for libraries.

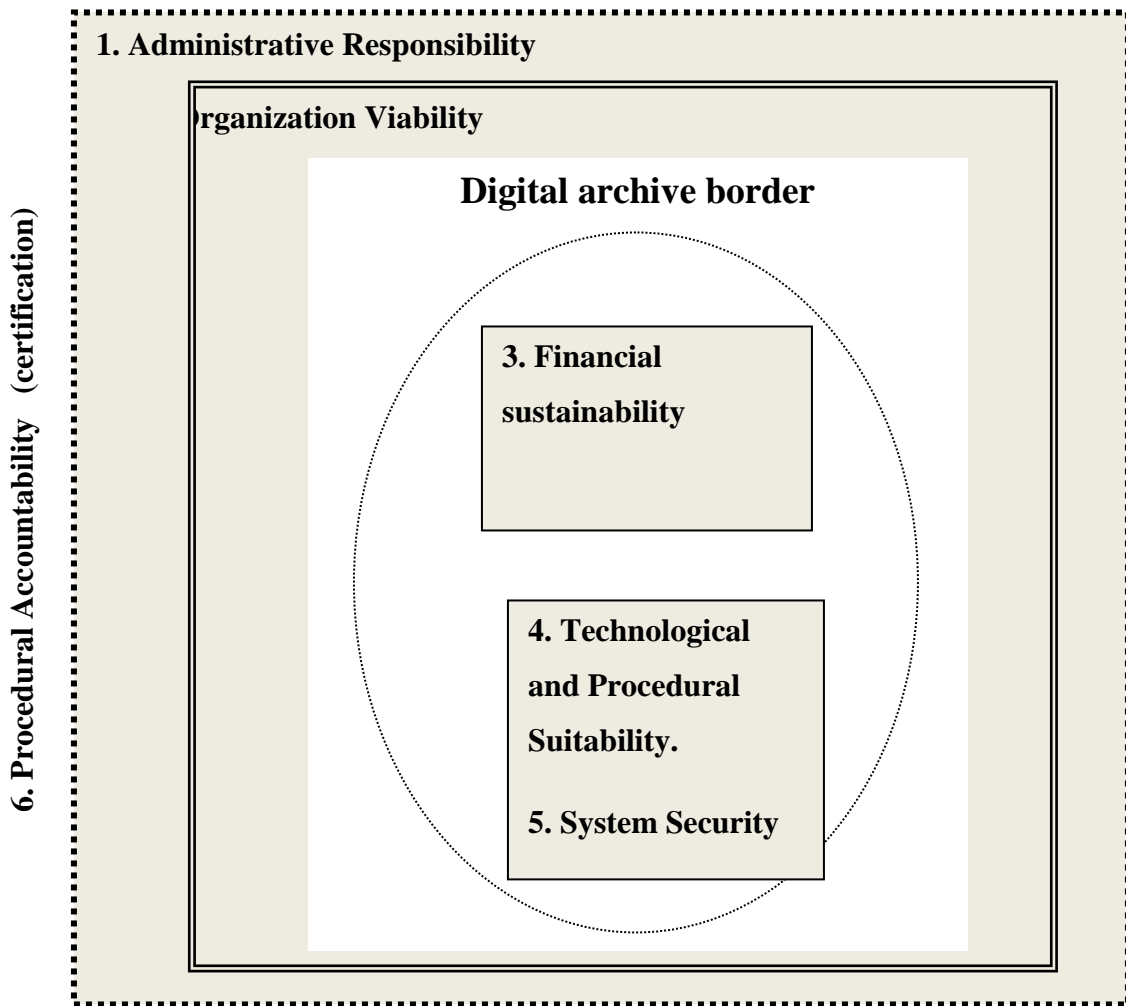
The ISO 16363:2011 advocates that an institution establishing a digital repository with preservation as a goal considers all preservation strategies available and identify

those that fit it. This ensures that it has in place all appropriate hardware and software needed and has plans on how they will be replaced. The standard further goes to recommend that the repository should comply with all relevant standards and best practices, ensuring that staff have adequate expertise to understand and implement them. The trusted digital repository should undergo regular external audits on its system components and performance. According to Interuniversity Consortium for Political and Social Research(ICPSR) (2009), technological suitability refers to existence of suitable and evidence of software, hardware, and the skills to establish and maintain the digital preservation programme and at the same time ability of the organisation to anticipate and respond wisely to changing technology. A trustworthy digital repository should undertake technology watch, plan for finances to replace software and hardware when obsolete, hardware and software capabilities to support backups, mechanisms to identify bit corruption and documentation capturing risks and loss and the strategies to recover loss (Dollar & Ashley, 2020). Procedural suitability on the other hand takes into account the existence of policies, procedures and best practices all tailored to addressing digital preservation. Lin et al (2020) argues that digital repositories can demonstrate technology suitability by implementing the relevant and appropriate standards, tools, and technologies for data management and curation as well as having plans and mechanisms to support them

In order to demonstrate procedural responsibility, a digital repository must assure stakeholders by documenting all processes as well as decisions and goals it has adopted (Bantin, 2016). In addition, a policy should be available indicating the

designated community, preservation policies in place, all preservation actions undertaken as well as hardware and software available (Bantin, 2016)

Lastly, trusted repositories must ensure that all repository practices are well documented and made available on request. The repositories must also have monitoring mechanisms that measure and ensure the continued operation of all systems and procedures are in place. All preservation actions undertaken should be recorded and justified in the context of community-wide best practices with feedback mechanisms in place to support the resolution of problems and to negotiate the evolving requirements between the repository, any third-party service providers, and the designated communities. Lin, et al (2020) insists that trustworthiness is demonstrated through evidence making it a requirement for digital repositories to provide transparent, honest, and verifiable evidence of their practices. Milam (2014), proposes some of these procedures to include but not limited to: procedures to prevent, discover and correct loss or corruption, procedures against media deterioration and technology obsolescence, establishment of documentary forms for each procedure, rules for authentication; procedures for identifying authoritative records, and procedures on the removal and transfer of relevant documentation. Figure 2 summarizes the attributes of a trusted digital repository as presented by ISO 16363:2011



*Figure 2: Trusted Digital Repositories Framework*

#### 2.1.2.1 Applicability of ISO 16363:2012 to the Study

Lin et al (2020) recognized the need for best practice in digital preservation but noted that the adoption of best practices was not enough, emphasizing that digital repositories had a duty to their users to demonstrate that they are reliable and capable of appropriately managing the information that they held. According to Wang (2017) institutional repositories needed to perform evaluations of these practices and proposed the TRD framework as a tool that can be used to perform the evaluation. This study aimed at establishing the efficacy of digital preservation practices in

selected public universities in Kenya and was guided by the following objectives: determine the characteristics of digital resources in the institutional repositories; audit the existing digital preservation policies; establish the existence of digital preservation planning; audit the digital preservation strategies practiced by the IRs in the selected universities; Identify digital preservation challenges in the selected public universities; and to make recommendations and propose best practices for digital preservation in The IRs of the selected universities.

ISO 16363:2012 presented a number of functions an institutional repository must enact in order to be trusted. These functions resonate well with the objectives of the study by enabling the researcher to identify the type of content ingested into the IRs and whether content was defined by policies and procedures as well as whether the IRs clearly followed the policies and procedures. The framework assisted the researcher to identify evidences that digital preservation was being carried out by acknowledging the role documentation of actions play in digital preservation. Other actions informed by the framework include, preservation planning, technology watch, budgeting, metadata creation and management among others.

**Table 2: Linking the Standards to the Research Questions**

<b>Research Question</b>	<b>ISO 14721:2012</b>	<b>ISO 16363:2012</b>
1. What are the characteristics of digital objects are accepted for storage in IRs in the public universities under study?	OAIS requires institutional repositories to negotiate for relevant content by defining the type of data (both in content and file format) accepted in the repository as well as providing procedures for those uploading materials.	Stipulates that Digital repositories should be compliant with the OAIS framework. Organisations should have procedures and policies
2. How suitable are the digital preservation policies to support digital preservation?	emphasizes the development of policies that support preservation and long-term access	Stipulates that organisations should demonstrate viability through their mission and goals as well as by having policies and procedures
3. How adequate are the digital preservation plans of the selected IRs?	Planning is one of the functions recommended by ISO 14721 and involves identification of preservation strategies, technology planning as well as taking part in research and development of digital preservation strategies	The ISO 16363 requires that trusted digital repository should be able to prove its financial sustainability over time and adhere to good business practices including a business plan and adequate budget and reserves.
4. To what extent do the preservation strategies being used by the IRs adequate long-term preservation?	recommends that digital repositories should adopt digital preservation strategies that will ensure long-term access to information resources by the designated community.	ISO 16363 advocates that a repository an institution establishing a digital repository with preservation as a goal considers all preservation strategies available and identify those that fit it. In addition, organisations should demonstrate organizational responsibility by frequently carrying out quality assessments and communicating the same to the customers

## 2.2 Need for Digital Preservation

The existing body of literature consistently underscores a critical distinction between the preservation of print resources, often referred to as traditional preservation, and the preservation of digital resources. This distinction arises from the fundamental nature of these materials. Traditional preservation practices were primarily designed to maintain the physical integrity of print materials, a relatively straightforward task compared to the challenges posed by digital preservation. Unlike print, where the objective was often to keep items unchanged, applying this same principle to digital objects would inevitably lead to formidable obstacles in ensuring future access to these materials (Gaur & Tripathi, 2012).

Pennock (2006) succinctly captures a critical concern in the realm of digital preservation when asserting that "Digital objects will cease to be accessible without active management and intervention" (p.1). This statement serves as a stark reminder of the ever-evolving nature of digital content and the pressing requirement for continual vigilance in its preservation. As Dharini (2009) aptly highlights, the responsibility for digital preservation rests not only with organizations but also with individuals. Effectively safeguarding digital resources demands a deliberate and collective commitment from all stakeholders involved in the lifecycle of these resources.

In this context, institutions, be they museums, archives, or libraries, find themselves facing a pressing need to implement short- to medium-term digital preservation programs. These initiatives are not only a matter of best practice but also crucial for legal compliance, particularly in the context of fulfilling access to information laws (Baucom, 2019). Therefore, the responsibility for digital preservation extends far beyond mere

technological considerations; it underscores a broader societal obligation to ensure the accessibility and longevity of our digital heritage.

Institutions engage in digital preservation for a number of reasons such as: usability (ensures that the intellectual content of the item remains usable irrespective of changes to technology); authenticity (the provenance of the content must be proven and the content an authentic replica of the original); discoverability (the content must have logical bibliographic metadata so that the content can be found by end-users through time; and accessibility (the content must be available for use to the appropriate community), (Patel, 2014). There are instances where the goals are not clearly defined but revolve around user expectations (Whitt, 2017). According to Traczyk, et al (eds) (2017) digital preservation should aim at “keeping digital materials not only technically accessible, but also usable for long periods of time” (p. 13). While statements emphasize important goals of digital preservation, they oversimplify the complexity of the preservation process. The focus on usability, authenticity, discoverability, and accessibility is undoubtedly crucial, but does not fully capture the multifaceted challenges and considerations involved in preserving digital materials.

From a researcher’s digital preservation should meet the following goals (Ross & Hedstrom, 2005): *The protection and conservation of cultural memory*: Hardin (2008) argued that universities while addressing their current problems needed to focus their responsibilities for cultural stewardship, suggesting that part of the stewardship of materials probably ought to be getting the digital surrogates in case of disaster overtaking the originals, thus promoting the preservation of rare and fragile objects without denying access to those who wish to use them. This further was emphasized by the Library of



*Congress's National Digital Information Infrastructure and Preservation Program (NDIIPP) 2010 report on Preserving our digital that concluded that preservation of digital information is a societal good that organisations must make effort to protect; Long term access to digital materials enables cross disciplinary collaborations:* International scientific collaborations have benefited from the availability of data repositories. In the perspective of Chitez et al. (2020), the concept of multi-disciplinary digital repositories takes on a pivotal role in fostering cross-disciplinary collaborations and enhancing international networking in groundbreaking research domains. It is evident that the scientific community's collaborative endeavors increasingly rely not only on their own datasets but also on the contributions of fellow colleagues, as highlighted by Curdt (2019). This underscores the pressing necessity for robust infrastructures capable of accommodating cross-disciplinary demands.

Furthermore, the imperative of digital preservation becomes apparent when considering the enablement of material reuse, a notion emphasized by Bote and Termens (2019). Reusing information, especially research data, empowers scholars to conduct secondary evaluations, reframe research inquiries, and facilitate longitudinal studies of temporal changes. As elucidated by Hedstrom et al. (2003), reuse stands as a primary economic benefit of digital preservation, as it promotes data sharing, provided the data is reliable, accessible, discoverable, and persistently stored, as articulated by Austin et al. (2016). In addition, researchers necessitate a pre-established assurance that their data will be maintained over an extended period, not only to facilitate reuse but also for citation purposes, in alignment with the insights of Bote and Termens (2019).

Moreover, the transition from industrial to knowledge economies is intricately linked to affordable and effective digital preservation strategies. A knowledge economy, as described by Chen & Qu (2008), places knowledge and information as the foundational drivers of economic growth, surpassing traditional factors of production such as capital and land. In this context, information and communication technology (ICT) emerges as a pivotal catalyst for the knowledge economy. Contemporary knowledge is predominantly stored electronically, encompassing files, databases, web content, and software programs, as noted by Jharota (2018). It is incumbent upon us to preserve this wealth of knowledge, ensuring its perpetual accessibility and legibility, transcending technological shifts and evolutions.

In line with the insights of authors such as Micunovic et al. (2016), it becomes evident that safeguarding digital resources from the perils of technological obsolescence is imperative if we are to fully reap the rewards bestowed by the digital age. Deridder (2016) succinctly encapsulates the lifespan of a digital object as intricately entwined with the formats employed, the requisite software and hardware for access, and the quality of content storage.

Indeed, the essence of digital preservation, as articulated by et al. (2015), transcends the mere process of safeguarding information; it hinges on the value derived from the activities conducted with the preserved object. Baucom (2019) underscores the universal obligation for organizations to preserve their digital content, particularly considering instances where digital surrogates are lacking. Long-term digital preservation stands as the paramount means to retain the informational assets that organizations have invested substantial time, financial resources, and personnel efforts in creating.

The sagacious counsel provided by Dollar and Ashley (2020) serves as a stark reminder to organizations embarking on the journey of ensuring long-term access. Their admonition is clear: digital preservation is a matter not to be taken lightly, as it underpins the sustained availability and utility of valuable information resources.

### **2.3 Digital Preservation and Institutional Repositories**

The discourse surrounding the involvement of Institutional Repositories (IRs) in digital preservation has been ongoing, marked by divergent viewpoints. As Hockx-Yu (2006) notes, there exists a divide between those advocating for digital preservation as an inherent function of the repository and those emphasizing the improvement of access, usage, and impact. Importantly, these debates do not negate the essential role of digital preservation within IRs; rather, its pertinence is contingent upon the specific goals and content of the institutional repository, a perspective supported by Thibodaux (2007). Thibodaux suggests that the criteria for measuring the success of an IR should be derived from its stated purpose, emphasizing that the mission statement should guide whether the repository engages in long-term or short-term digital preservation.

In essence, Dowding (2016) underscores the intrinsic connection between access and digital preservation, positing that one cannot exist without the other—a symbiotic relationship. IRs serve as repositories for the scholarly output of universities, and as Hockx-Yu (2006) aptly notes, researchers, students, staff, and institutions alike rely on the ongoing availability and future accessibility of repository content. This responsibility places a substantial burden on IR administrators, who must ensure the enduring availability of resources, which could span decades or more. Knight (2005) echoes this

sentiment, emphasizing the need for digital preservation to guarantee the accessibility and integrity of academic research stored within repositories over the long term.

However, IRs have faced criticism for not fully embodying the characteristics of dedicated digital preservation repositories, leading them to prioritize access over preservation, as observed by the Center for Research Libraries (CRL) in 2008. This, in part, is attributed to IRs failing to explicitly make long-term access to digital resources their mission for designated communities (Yakel et al., 2008). Dell and Shultz (2014) emphasize the pivotal role of mission-driven commitment, as it directly impacts future access to repository resources. Corrado and Sandy (2017) contend that while IRs excel in creating infrastructure and providing access to institutional digital content, they often fall short in delivering the active management necessary for long-term preservation. Dowding (2016) points out that IR managers sometimes expect these repositories to preserve digital resources indefinitely without providing a comprehensive understanding of the expectations. It's worth noting that despite IRs not always adopting a forward-looking perspective, they still house unique digital objects not found elsewhere, further underscoring the necessity of digital preservation, as highlighted by Frank (2018) and Francke et al. (2017).

Alien (2006) outlines a comprehensive framework comprising four essential pillars for the effective execution of digital preservation programs. These pillars encompass the critical requirement for garnering unwavering support and commitment from stakeholders, the implementation of robust best practices, the cultivation of collaborative relationships and partnerships, and the establishment of comprehensive policies, standards, and procedures. Building on this foundation, Baucom (2019) underscores the

multifaceted nature of success in digital preservation, emphasizing the indispensability of various key partners. These encompass a diverse spectrum, encompassing resource allocators, institutional leadership, content creators, internal collaborators, external collaborators, and the end users, all of whom play crucial roles in bolstering the program's efficacy. External collaborators can even extend to include marketing experts and commercial vendors specializing in digital preservation tools, as elucidated by Barons et al. (2021). It is worth noting that these categories are not mutually exclusive, as individuals or entities may assume multiple roles, and the involvement of different stakeholders can vary at different stages in the lifecycle of a digital object, a concept illuminated by Ravenwood et al. (2015). Therefore, the continuous communication and clarification of the responsibilities and contributions of each stakeholder becomes paramount, as emphasized by Keller and Cordeiro (2020). Among these stakeholders, users occupy a pivotal position, as the selection and preservation of content should align seamlessly with their needs and expectations, a perspective eloquently articulated by Sandy and Corrado (2017).

## **2.4 Digital Preservation Practices**

The free online dictionary defines practices as the habitual or customary action or way of doing something. ISO 14721, 2012: OAIS and ISO 16363/TRD, 2011 identify criteria by which effective digital preservation in digital repositories can be evaluated. Included in the criteria is appropriate content, policies and plans as well as the actions undertaken to perform actualize digital preservation

### **2.4.1 Content in Institutional Repositories**

Johnson (2002) defines Institutional Repositories (IRs) by four key attributes: institutional definition, scholarly nature, cumulativeness, and openness with interoperability. The institutional definition implies that IR content originates exclusively from one institution, as Genoni (2004) points out. Researchers have characterized the content of digital repositories as encompassing various materials, including peer-reviewed journal articles, theses, dissertations, research data, monographs, book chapters, conference proceedings, departmental newsletters, grant-related papers, and reports to funding agencies (Genoni, 2004; Waddington et al., 2013). Two factors contribute to this diverse collection scope. Firstly, as highlighted by Shearer (2006), the term "scholarly" doesn't restrict repositories from accommodating other materials such as university annual reports, videos, computer programs, datasets, and photographs. Secondly, Breytenbach et al. (2013) advocate for IRs as guardians of information permanence and accessibility, promoting the inclusion of not just scholarly output but also unpublished conference papers, teaching materials, unpublished research, and corporate materials like institutional publicity.

The incorporation of peer-reviewed journal articles into IRs can be traced back to the Open Society's 2002 declaration that open access peer-reviewed journal literature was the way forward for scholarly communication (Budapest Open Access Initiative 2001). This led researchers to publish their findings through green and gold open access journals. While some peer-reviewed journal articles may have analog equivalents, others are born digital and pose digital preservation challenges. Rieger et al. (2022) observed that institutions tended to prioritize the preservation of digitized content that could be re-

digitized, neglecting robust strategies for the long-term management of born-digital content, which, if lost, might be irreplaceable.

Research data, comprising underlying or raw data supporting publications, has found its way into institutional repositories due to mandates from research funders, particularly in the UK, where grant holders are compelled to make their data open access post-analysis and publication (Pryor and Donnelley, 2009). Entities like the UK's Economic and Social Research Council (ESRC) enforce data sharing, enabling cost-effective data reuse for result verification, comparison, and generation of new knowledge (Mauthner, 2013).

The diverse collections within IR content demand varying digital preservation approaches (Hockx & Brower, 2018). While some collections have analog equivalents, others are entirely digital. Hockx and Bower (2018) argue that not all content within digital repositories requires long-term preservation, necessitating typologies (item classifications) to delineate the scope of digital preservation. Smith (2008) recommends a needs assessment to ensure that ingested content meets the designated community's requirements. Selection proves vital in the digital preservation process, supported by Lunghi et al. (2012) and Smallwood (2020), who underscore the importance of selecting and appraising digital materials as best practice for organizations committed to long-term digital preservation. Ismail and Affandy (2018) posit that resource constraints on those responsible for long-term preservation make selection inevitable, driven by the need to maintain the usability and searchability of the bitstream, necessitating specialized planning skills, time, and human effort, while avoiding potential damage to databases (Simon & Kiszl, 2021). Selection, which may appear negligible initially, becomes a

significant concern and resource burden as repositories grow, requiring additional space, IT expertise, and interventions to ensure content accessibility (Ravenwood et al., 2015).

A multi-stakeholder involvement and especially archivists in the selection of digital repository content has been recommended since they have expertise in appraisal that ensures the identification of resources that require long-term preservation (Dell & Shultz, 2014; Smith, 2008; Tallman & Work, 2018). Keller, et al (2019) in their report “Digital Preservation Task Force Update”, established that although archival expertise had been identified as critical success skills in the establishment of IRs, very few cases involved the archivist and in cases where they were involved, their expertise was visible in the scope and content of the institutional repository, adoption of preservation standards and the inclusion of structural and technical metadata for preservation purposes. According to Tallman & Work (2018), collection developers leave all preservation decisions to Preservation experts due to discomfort or unfamiliarity brought about by the technical nature of digital objects.

It is important to note the relationship of content to planning as summed up by Webb, et al (2013) who pointed out that when organisations clarified their preservation intentions it was a likely good starting point for preservation planning for diverse digital collections as it was adept to identify what needs to be kept and what does not warrant the use of limited preservation resources giving credence to the importance of selection and planning and the interconnectedness of a digital preservation activities.

Institutional repository administrators receive content is received either through a CD-ROM, email or links provided by the creators. This has an impact on the preservation



strategies adopted because decisions will have to be made on whether to preserve the link, email or storage media for born digital resources. Nadal (2007) recommended that there was need to ensure that these storage media were secure and reliable as they were known to be fragile and unstable.

#### **2.4.2 File Formats**

Brown (2006) describes file formats as the internal structure and encoding of digital objects that allow them to be processed or rendered in human-accessible forms. Digital preservation aims to ensure the long-term accessibility of digital objects but faces challenges stemming from technological obsolescence, which can render old file formats unreadable and unusable, as noted by Barve (2007). Various file formats cater to different applications, including text, audio, video, images, databases, presentations, spreadsheets, and markup languages. The choice of file formats and preservation media is critical. Lundell (2012) suggests that file formats for digital preservation should be hardware and software-independent, as the information they encode typically outlives the hardware and software used. Open file formats are recommended by Zuccala (2006) for longevity, protection, and preservation. Rimkus et al. (2014) propose the establishment of file format policies to guide repository managers, while Termens et al. (2015) highlight the lack of consideration for file format concerns as a preservation action, emphasizing the importance of quality control in file format selection. Francke et al. (2017) found that many repositories in Sweden had guidelines for accepted file formats but few considered whether these formats constituted open standards. They also noted that PDF files were often encrypted, potentially hindering future file migration, a crucial aspect of successful long-term preservation.

File formats can be classified into those recognized by standardization organizations and maintained by specific companies. For instance, PDF, developed by Adobe Systems Incorporated, was released as an open standard to the International Organization for Standardization (ISO), leading to the creation of PDF/A (archival). ISO defines PDF/A as a file format that ensures the visual appearance of digital objects remains consistent across different tools and systems, making it suitable for preserving textual data (Property Records Industry Association, 2017).

The Tagged Image File Format (TIFF) was created by the Aldus Corporation in 1986 for use in desktop publishing with the current TIFF 6.0 developed in 1992 and its rights acquired by Adobe Systems Incorporated in 1994. Fornaro, et al, (2017) have hailed TIFF as a format suitable for archiving because it is well documented, widely used and lacking proprietary elements. Table 3 below shows examples of recommended open file formats.

**Table 3: Recommended Open Standard Technology Neutral File Formats**

	PDF/A	XML	TIFF	PNG	JPEG 2000	SVG	MPEG-2	BWF	WARC
Text	✓	✓							
Spreadsheets	✓								
Raster images	✓		✓	✓					
Photographs					✓				
Vector graphics						✓			
Moving images							✓		
Audio								✓	
Web									✓
Databases		✓							

**Adopted from Dollar & Ashely, 2020**

File format registries play a crucial role in assisting institutions and organizations in making informed decisions about which file formats to use for long-term digital preservation. As described by Barve (2007), a file format registry serves as a repository of format specification information, encompassing descriptive, administrative, and technical metadata about digital formats. This information includes the syntactic and semantic characteristics of registered formats. Ryan (2014) succinctly encapsulates the purpose of file format registries, which is to gather and disseminate comprehensive information about file formats, including compatible rendering software, associated risks, and mitigation strategies in case formats become unrenderable. Notable examples of file format registries include PRONOM, established by The National Archives of the United Kingdom, providing a database of file formats with detailed characteristics and relationships between different formats. GDFR (Global Digital Format Registry), developed by the University of Harvard Digital Library in collaboration with Online Computer Library Centre (OCLC) and the US National Archives and Records Administration (NARA), and UDFR (Unified Digital Format Registry) are also among the most common registries, as noted by Shala and Shala (2016).

Despite the existence of these valuable resources, file format policies in Institutional Repositories (IRs) often suffer from non-compliance, as highlighted by Barrueco and Termens (2022) and Rimkus et al. (2014). Institutions frequently prioritize content creation over quality by accepting various file formats from creators without stringent adherence to established policies. This lax approach to file format policies has implications for long-term digital preservation efforts.

### **2.4.3 Digital Preservation Policy**

Shiloba and Mohammed (2013) define a digital preservation policy as a plan of action for the safekeeping of digital objects that addresses the questions of what needs to be preserved, why, for what purpose, and for how long. A digital preservation policy facilitates the effective management of the digital records ensuring the organisation is able to carry out its mandated functions, (International Research on Permanent Authentic Records in Electronic Systems (Inte rPARES) Project, 2008). Dell & Shultz (2014) noted that most organisations have the mistaken belief that digital preservation occurs without conscious intervention and note that the presence of a digital preservation policy shows that the organisation has made a conscious effort to digital preservation. They went on to add that institutional repositories required clear guidelines that defined what was to be collected and saved a view rightly supported by Emmott (2008) by declaring that without policies, practitioners have little to guide their decisions about what must, should, could and won't be preserved, let alone how. Furthermore, the lack of preservation policies showed a lack of commitment to digital preservation by institutions (Cloonan & Sanett, 2002).

Policies are vital for ensuring compliance with procedural and legal requirements within an organization by clearly defining a set of procedures, roles and responsibilities in order to promote accountability as well as ensuring that preservation activities are included in strategic plans (The British Library, 2013). Also, policies provide an overall cohesion within an organization and offer guidance for best practice (Dressler, 2017).

According to the UK National archives, (2011) a digital preservation policy has two overarching purposes that include defining why an organisation is doing digital

preservation and the benefits of long-term access to data for re-use and the economic, reputational and cultural risks of failing to address this through digital preservation a view shared earlier on by NISO (2010) by putting forward that the policies articulated clearly how digital assets will be managed in a repository to avert the risk of content loss by specifying amongst other things, data storage requirements, preservation actions, and responsibilities. Earlier on Jones and Semple (2006) rightly pointed out that organisations needed to undertake digital preservation needs as a cross-disciplinary responsibility since relevant skills needed may be spread throughout an organisation. In this regard then, policy acts as the authority for those undertaking digital preservation in terms of not only articulating roles and responsibilities both within the organisation and any external parties (contractors, depositors / donors of records) but also with creators of digital materials who need to be able to understand the implications of their actions on the medium to long-term viability of the digital material they create (DPC, 2015). Policies act as guidelines and communicate the services and actions that should take place routinely (Dressler, 2017).

Ismail and Affandy (2018) described policy development as one of the most vital digital preservation strategies as it stipulated what needed to be preserved, for what purpose and for how long. Despite this, researchers have consistently found a lack of digital preservation policies within institutions with studies by Waller & Sharpe, (2006) and Beagrie, et al (2008) highlighting the lack of digital preservation policies in libraries. Sinclair et al (2011) citing a survey by the Planets project (Preservation and Long-term Access through Networked Services) indicated that about 43% of libraries who responded had a digital preservation policy and that those who had the policies were able

to budget for it to even include it in their organizational planning and more so had solutions for it. In fact, Roy, et al (2018) reported IRs all over the world as lagging behind in preservation policy development only defining the file formats likely to support long term preservation.

Beagrie, et al(2008) summed it up very well when they argued that the lack of a policy leads to the lack of consideration of digital preservation issues in other institutional strategies and termed this a major drawback to preserving digital information a point that Coonan and Sannet (2002) had earlier on pointed out that preservation policies indicated that the organisation was accountable and increased trust that funding will be utilized for the long term good of the organisation. They were also of the view that preservation procedures could only be implemented wholly with adequate funding a view emphasized by the DPC (2015) who added that a policy could be used to seek funding. Beagrie et al (2008), emphasized the need to integrate the preservation into business drivers, activities and functions such as regulatory compliance, staff development, applied technology and academic excellence.

Policy statements set clear priorities, and ensures stakeholder collaborations making it the responsibility of the organisation safe guard its digital resources for the future generations and ensured that it also defines organizational roles and (funding) responsibilities. Ismail and Affandy (2018) citing the National archives of Georgia (2007) recommends the inclusion of collaborations and partnerships as part of digital preservation strategies in several areas such as working with government agencies to developing legislation and procedures and in developing policies that support long term preservation with creators and users.

The need to frequently review digital preservation policies cannot be overstated. Cloonan & Sannet (2002) acknowledge that policies need to evolve rapidly since the policy owners may encounter new and unanticipated features that require new policy decisions. In fact, they should be updated routinely to ensure their validity and usefulness (Dressler, 2017). Although Madsen and Hurst (2019) advocate for a specific shelf life for the policy (between 3-5 years), literature shows variations in digital preservation policy review cycle among organisations with some even lacking the same in their policies. For example, commenting on a digital preservation task force formed by the Ohio State University Libraries, Noonan (2014) observed that of all the digital preservation policies evaluated by the task force, only one had a statement to maintain the currency of the digital preservation policy through regular reviewing. Friese (2012) when developing digital preservation guidelines for the NESTOR group noted that Marriot Library of the University of Utah in Salt Lake City, USA, had revised its policy three times during the last three years to take into consideration the rapid changes in technology. She goes on to argue that when an organisation revises its policy regularly, it shows that it actively watches technology and developments in digital preservation and keeps the preservation policy up to date. The UK National archives (2011) in their *guidelines to the development of digital preservation policies: guidance for archives* recommends a period not exceeding three years to factor in technological and contextual changes while Brown (2013) felt that a period of 2 years was adequate to warrant an update of the policy.

Barrueco and Termens (2022) in a review of literature on digital preservation policies noted that although IRs in Africa had policy statements, these could not be equated to a proper policy. Instead the statements only expressed the commitment to preserve

contents but the literature lacks evidence on the existence of digital preservation policies within IRs in Africa.

#### **2.4.4 Digital Preservation Planning**

Becker et al. (2009) offer a comprehensive understanding of preservation planning as a sequence of actions intended to mitigate risks and ensure the long-term access and authenticity of digital objects within a designated collection. Kool et al. (2014) on the other hand differentiate preservation policies from plans, emphasizing that policies provide overarching guidance and frameworks, while plans provide actionable steps for achieving long-term access. Dressler (2017) astutely underscores the interconnectedness of policy and plans, highlighting that a well-crafted policy without a corresponding strategic action plan remains a mere document, and conversely, actions without strategic alignment lack documentation and a clear conceptual foundation. Preservation planning is recognized as a foundational capability in digital preservation, forming the core of the OAIS model (ISO 14721, 2012), as acknowledged by Dollar and Ashley (2015). It is instrumental in guiding decision-making, assessing preservation needs, and delineating the workflow for evaluating and defining preservation plans within an organizational context (Becker et al., 2014).

Digital preservation planning is not merely a theoretical exercise but has practical implications for organizations, as emphasized by Becker and Rauber (2011). It supports decision-making at various management levels by establishing criteria for preservation, defining workflows, and identifying risks that may impact long-term access to digital resources. Yannis and Yannis (2018) emphasize the need to identify specific preservation needs, list available solutions, and define actionable steps, all of which contribute to the



development of a comprehensive preservation policy framework. Organizational preparedness for digital preservation is deemed critical, considering the higher costs associated with preserving digital materials compared to traditional paper collections, as emphasized by the DPC (2008).

The benefits of digital preservation planning, as outlined by Becker et al. (2009), are extensive and encompass enabling preservation actions, documenting those actions, ensuring accountability, and facilitating consistent and ongoing management of digital objects. Financial sustainability is a crucial aspect of digital preservation planning, aligning with the TDR's emphasis on financial sustainability. Rinehart et al. (2014) advocate for incorporating digital preservation into existing plans to secure the necessary financial resources, especially given the resource constraints and the evolving nature of technology in the digital preservation landscape, as underscored by Harvey (2005).

However, the primary challenge in integrating digital preservation into organizational processes lies in the general lack of awareness among stakeholders, as noted by the International Records Management Trust (2016). This results in the omission of digital preservation from planning processes. To address this, the trust suggests that successful digital preservation planning should be integrated with an organization's broader strategic directions and priorities.

#### **2.4.5 Metadata and Digital Preservation**

Preservation metadata encompasses critical information required for ensuring the long-term accessibility of digital objects. Given the inherent reliance of digital resources on specific computing environments and the rapid pace at which technology becomes

obsolete, it becomes imperative to preserve not only the digital resource but also accompanying details about the hardware, software, file formats, and storage media in which they were originally created. Dappert et al. (2016) emphasize that merely storing digital objects on a data carrier is insufficient. Digital repositories must adopt strategies that not only shield digital objects from accidental or intentional damage but also facilitate the recreation of the complete computing environment, if the need arises. This becomes crucial because alterations in the computing environment can lead to changes in the rendering and presentation of the digital resource, as highlighted by Zuccala (2006).

Besser (2000) delineates preservation metadata as a strategic approach aimed at furnishing comprehensive technical information about resources while supporting migration and emulation as preservation strategies. By documenting the technical environment of a resource, preservation metadata enables effective migration to newer hardware, software, storage media, or file formats. Furthermore, it allows for the faithful emulation of the previous environment when necessary, a concept articulated by Woodyard (2004). Woodyard underscores that preservation metadata should encompass the ability to list technical details about files and resource structure, document the history of all actions performed on the resource, authenticate its integrity through technical means, and maintain records about custody, responsibility, and rights associated with preservation actions.

Cendi (2006) aptly underscores the need for all organizations to document their purposes and requirements, including those related to preservation metadata. Jones (2006) concurs with this perspective and extends it by advocating not only for substantive descriptive metadata to enhance access but also for technical metadata to facilitate preservation.

ISO 14721 (2012) defines Preservation Description Information (PDI), which includes reference, provenance, context, fixity, and rights metadata, as outlined by Giaretta (2011). Several scholars emphasize the significance of metadata in digital preservation, with RLG (2002) emphasizing that the creation and deployment of preservation metadata are pivotal components of most digital preservation strategies. Preservation metadata has been acknowledged as a best practice for the long-term stewardship of digital resources (DPC, 2006; Gartner & Lavoie, 2013; Joo et al., 2019). Nevertheless, Barrueco & Termenes (2022) note limited evidence of preservation metadata usage, with most studies predominantly focusing on descriptive metadata, as observed in research on ARL libraries.

Metadata standards play a pivotal role in gathering the appropriate metadata. Standards like Dublin Core, METS (Metadata Encoding and Transmission Standard), and PREMIS (Preservation Metadata Implementation Strategies) define the creation and management of metadata (El Idrissi, 2019). However, IRs encounter challenges in selecting the most suitable standard, given the diverse services metadata must support, including access, preservation, and ingest, as observed by Joo et al. (2019). The suitability of Dublin Core for preservation metadata is a subject of debate, with Joo et al. (2019) highlighting its limited vocabulary, while Mukherjee and Das (2020) contend that, although inadequate as a metadata schema, it offers flexibility and can be customized to accommodate more comprehensive metadata schemas.

#### **2.4.6 Digital Preservation Strategies**

The UK National Archives (2011) describes a digital preservation strategy as a broad approach adopted by an organisation to ensure that the content of digital records remains

in a usable form over time. There is as yet no single strategy that can achieve long-term preservation and access to all the types of digital resources stored in digital repositories. Bountour i(2017) proposes that strategies should be based on proactive preservation indicating that these strategies should be applicable throughout the life of the digital object. According to Barrueco & Termenes (2022) digital preservation strategies should start at the creation of the digital objects by raising awareness among the creators. UNESCO's Guidelines for the Preservation of Digital Heritage (2003) group digital preservation strategies into four groups:

#### **2.4.6.1 Short-term Strategies**

These are digital preservation strategies that are geared towards providing access to digital resources for a predetermined period of time. They include:

**Bitstream Copying:** This is the process of making an exact duplicate of a digital object and is also known as backup. It works in combination of other digital preservation strategies like remote storage, refreshing and migration. Backups are designed to protect data in the event of hardware failure or other catastrophic events, but they are not intended to be a long-term preservation solution (Arora, 2009).

**Refreshing:** refers to copying of digital information from one long-term storage medium to another of the same type, with no change whatsoever in the bitstream (e.g. from a decaying 4mm DAT tape to a new 4mm DAT tape, or from an older CD-RW to a new CD-RW). Refreshing is a necessary component of any successful digital preservation program, but is not itself a complete program. It potentially addresses both decay and obsolescence issues related to only the storage media. Lee, et al (2002) echoed this clearly by stating that refreshing involves copying the digital information onto newer

media before the old media becomes so obsolete that the data cannot be accessed and that this required predicting the lifetime of a digital media. They went on to add that refreshing depended on well preserved bits further supporting the view that no single strategy could work on its own.

Technology Preservation: this strategy involves preserving the computing environment and it is sometimes called the "computer museum" solution. It has been argued that technology preservation is more of a disaster recovery strategy for use on digital objects that have not been subject to a proper digital preservation strategy (Arora, 2009; Ismail & Affandy, 2018). Others have argued that the strategy offers the potential of coping with media obsolescence, assuming the media hasn't decayed beyond readability (Arora, 2009; Digital preservation Management, 2009; Harvey, 2011). They go on to add that although it can extend the window of access for obsolete media and file formats, it is ultimately a dead end, since no obsolete technology can be kept functional indefinitely and it is a strategy that no individual institution can implement, indicating the need for more than one institution to join together and implement it (Arora, 2009; Digital preservation Management, 2009; Harvey, 2011). This resonates with Lee, et al (2002) who affirmed that maintaining obsolete technology in usable form required a considerable investment in equipment and personnel. They also warned the strategy should be taken as a short term strategy because with age equipment breaks down, vendor support stops, documentation disappears. Feeney (1999) was of the opinion that managers who relied only upon this approach would may end up with "a museum of ageing and incompatible computer hardware"

Digital Archaeology: According to Lee et al (2002) digital archaeology includes methods and procedures to rescue content from damaged media or from obsolete or damaged hardware and software environments. It has also been described as an emergency recovery strategy that involves specialized techniques to recover bit streams from media that has been rendered unreadable, either due to physical damage or hardware failure such as head crashes or magnetic tape crinkling (Arora, 2009; Mohanty & Das, 2014; Whitt, 2017). According to Arora (2009), digital archaeology is mostly done by for-profit data recovery companies that maintain a variety of storage hardware (including obsolete types) plus special facilities such as clean rooms for dismantling hard disk drives.

Backward Compatibility and Version Migration: Backward Compatibility relies on the ability of current versions of software to interpret and present digital material created with previous versions of the same software and to save them in current format as evidenced by the latest versions of most popular word processing packages, for example, being able to decode files created on earlier versions of the same package (Lee, et al, 2002; Verdegem & Slats, 2004). Several researchers have however warned that it would be naïve for an organisation to rely on this as a long term digital preservation strategy or for more complex digital resources (Lee, et al 2002; Borghoff et al. 2006; Arora, 2009) while (Verdegem & Slats, 2004) recommend that records maintained this way be resaved in new file formats as it can only support a number of older generation file formats.

Version Migration involves permanently converting documents into a format that can be presented by the current version of the software (Najar & Wani, 2019). Thibodeau (2002) postulates that it involves migration from same family of products or data types. He goes on to add that successive versions of given formats provide a defined linear migration

paths for files stored in the format and that software vendors supply conversion routines that enable newer versions of their product to read older versions of the data format and save them in the current version. He however cautions against relying on version migration as a long-term digital preservation strategy as with time the format may become obsolete and also may present the preserved documents with characteristics they did not, and perhaps could not, have had affecting the authenticity of the documents.

Replication: Replication is a process whereby identical, multiple copies of files, file systems, or websites are created. According to Luan and Nygard (2010), it is an approach that involves the preservation of information resources using several copies and in different formats saved at different places. Kolle, et al (2014) argue that data that exists as a single copy in only one location is highly vulnerable to software or hardware failure, intentional or accidental alteration, and environmental catastrophes such as fire and flooding and that it is more likely to survive if it is replicated in several locations. They go on to add that due to the existence of the data in multiple locations, it may be difficult to implement other preservation strategies a notion Kolle, et al (2014) agree with terming it a simple but expensive strategy.

Although short term digital preservation strategies are easier to implement, they are not sufficient to guarantee long term access and integrity of digital resources but rather they are more useful for risk recovery and business continuity (Zuccala, 2006). In addition, short term strategies may result in large quantities of digitized objects or specialized tools that stagnate and become obsolete in time (Baucom, 2019).

#### **2.4.6.2 Medium- to Long-term Strategies**

According to Arora (2009) medium to long-term digital preservation strategies are aimed at providing continued access to digital materials beyond changes in technology for a defined period of time but not indefinitely and include the following:

**Durable/Persistent Media**(e.g., Gold CDs):This involves the use of digital storage media with a long life span and higher durability. Deshpande, (2016) argues that the strategy reduces the need for refreshing, and helps diminish losses from media deterioration, as do careful handling, controlled temperature and humidity, and proper storage but cannot mitigate against loss of information through physical loss, media obsolescence, as well as obsolescence of encoding and formatting schemes. At the same time, it gives the organisation a false sense of security making it to relax its preservation efforts and this way endangering the content. Harvey (2005), terms the use of durable media as a non-solution to digital preservation since does not address technological obsolescence but rather helps organisations to buy time while developing and implementing other viable strategies and practices that will make long-term preservation possible.

**Migration:** Migration refers to copying data, or converting data, from one technology to another, whether hardware or software, preserving the essential characteristics of the data. It encompasses a set of organized tasks designed to achieve the periodic transfer of digital materials from one hardware/software configuration to another, or from one generation of computer technology to a subsequent generation. Barateiro, et al (2009) summed this up by describing migration as any process that includes, conversion of digital information to analogue formats, version updates, conversion to other formats and normalization. Research puts the “purpose of migration as to preserve the integrity of



digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology” (Barateiro, et al, 2009; Harvey & Mahard, 2013). Migration includes refreshing as a means of digital preservation but differs from it in the sense that it is not always possible to make an exact digital copy or replica of a data base or other information object as hardware and software change and still maintain the compatibility of the object with the new generation of technology. Migration theoretically goes beyond addressing viability by including the conversion of data to avoid obsolescence not only of the physical storage medium, but of the encoding and format of the data.

Migration also includes changing from one file format to another sometimes referred to as forward migration. Caplan (2010) describes forward migration as “the process of creating a derivative version of a source file in a format considered to be a successor to the original format, either directly” (p. 226). Two surveys done in 2004, the first on 21 natural science and scientific publishing organizations operating on an international level and the other by OCLC on digital preservation practice in national libraries, state libraries, university and research libraries and consortia, archives, museums, and other organizations from 13 countries identified migration as the most common strategy utilized for digital preservation, although it was done in combination with other strategies (Harvey & Mahard, 2009). Digital repositories need to be cautious when carrying out format migration because it may lead to other risks such as of loss of content or context as well as inaccessibility of the information (National Archives of the UK, 2011).

Emulation: Emulation combines software and hardware to reproduce in all essential characteristics the performance of another computer of a different design, allowing programs or media designed for a particular environment to operate in a different, usually newer environment. According to the National archives of the United Kingdom (2007), emulation is meant to overcome technological obsolescence and involves reproducing the same systems or technologies for digital records activities by using the current generations of technologies and at the same time ensuring that the whole content, context and structure of the digital records will never change. Ismail and Affandy (2018), believed that the emulation maintained the exact look of the records that need to be preserved endeared it to many digital preservers keen on maintaining the originality of digital records.

Canonicalization: Arora (2009) describe canonicalization as a “technique designed to allow determination of whether the essential characteristics of a document have remained intact through a conversion from one format to another” (p. 126). Lynch (1999) asserts that canonicalization should be able to support preservation strategies like migration by ensuring that significant properties of the digital object are captured. Canonicalization works on the premise that a canonical form capturing the essential characteristics of a class of objects can be defined and that it is possible to translate data formats used to encode a given type of object to the canonical form (Ayre and Muir, 2004). Its greatest weakness is that because the canonical form does not include irrelevant data, reversing the translation may not give an identical object (Ayre and Muir, 2004; Najjar&Wani, 2019).

### **2.4.6.3 Investment Strategies**

Arora (2009) describe investment preservation strategies as involving an investment of efforts at the time of archiving digital materials. They include such strategies as: restricting formats and standards, reliance on standards, data abstraction and structuring, encapsulation, software re-engineering and universal virtual computer.

Reliance on standards: According to Gbaje (2011) the use of standards as a digital preservation strategy involves the adoption of “open, widely available, supported or agreed standards and file formats, for which there is an increased likelihood of stability and longer term support”. Strodl, et al (2007), give the adoption of PDF/A file format as a good example of reliance on standards for digital preservation an issue earlier on identified by Hedsrom (2001). Hedsrom (2001) had also argued that the digital preservation community tended to favor open standards proprietary standards since the open standards were published and were readily available unlike the proprietary standards which if adopted, their success was dependent upon the longevity of the firm that owns the standard or on its continued market dominance. A research done by Masenya and Ngulube (2019) identified the failure to adopt standards as one of the impediments to digital preservation. Hedsrom (2001) was quick to warn libraries against solely relying on standards as a preservation strategy suggesting that standards were useful if there was a consensus within the preservation community on which standards to adopt for the variety of resources that existed and they were readily available products that conformed to or supported the standard and if the standard had a demonstrated track record.

In addition to standards restricting formats through policies and procedures is also a long-term strategy. As Johnstone (2020) says, digital preservation adopts the life cycle concept and starts with the people creating the files, not when the files are ingested into the repository. This then requires “Guidance for Creators”

Normalization: CLIR and LoC (2002) describe normalization as the aspect of collecting various file formats and conversion them to one file format. In essence it is a formalized implementation of reliance on standards. It involves gathering all the digital objects of a particular type (for example color images, structured text) within a repository and converting them into a single chosen file format that is thought to embody the best overall compromise amongst characteristics such as functionality, longevity, and preservability (Arora, 2009). Caplan (2008) argued that normalization had the advantage of keeping the number of file formats the digital repository had to manage few.

Encapsulation: encapsulation has been described as “a technique of grouping together a digital object and metadata necessary to provide access to that object”(Beagrie, 2002, p.108). Ostensibly, the grouping process lessens the likelihood that any critical component necessary to decode and render a digital object will be lost. Appropriate types of metadata to encapsulate with a digital object include reference, representation, provenance, fixity and context information (Deshpande, 2016). According to Harvey (2005), encapsulation is a prerequisite to emulation.

Universal Virtual Computer (UVC):Lorie (2001) describes a UVC as a Computer in its functionality, Virtual as it will never have to be built physically and Universal because its definition is so basic that it will endure forever suggesting that all that was needed was to

write a UVC interpreter that could be written for any machine without changing the UVC program that is independent of architecture. According to Mohanty & Das (2014) a UVC requires the development of a computer program that is hardware and software independent and has the ability to simulate the basic computing environment architecture of every computer since the beginning. It uses both the elements of migration and emulation

#### **2.4.6.4 Alternative Strategies**

These include strategies such as use of analog backups and data recovery. Analog Backups: This strategy combines the conversion of digital objects into analog form with the use of durable analog media. Hoke (2012) give an example of converting digital to paper arguing that paper offered a longer lifespan but at the same time warned that the repository risked losing some digital qualities such as accessibility, lossless transferability and increase in costs especially storage if volume increases a feeling supported by Harvey (2005) who termed it as a non-solution.

Data recovery also called data archaeology is the recovery of information from damaged media and restoring the intelligibility of the data (UNESCO, 2003). It is a strategy that is used after data loss. It is however an expensive strategy that is done by profit making organisations. Magama (2017) argues organisations to develop quality preservation programs to avoid using this strategy.

#### **2.4.7 Repository platforms and digital preservation**

Management, content and technology are critical aspects in digital preservation (Corrado & Sandy, 2017). Management involves development of policies, plans and procedures

while technology involves hardware and software, file formats and storage media, tools and workflows, a secure environment, platforms and networks as well as the skills to establish and maintain the digital programme. According to Dollar & Ashley (2020) long term continuity of digital resources does not happen by accident but requires among other things planning, technology monitoring, adoption of evolving standards as well as monitoring of computing trends. Over time many projects have been undertaken to provide technological solutions that will support long term preservation of digital objects. Few offer full-service that includes digital storage, access, management and preservation with majority assisting with only individual aspects of digital preservation, such as ingest, access, or storage (McGuirk, 2016).

Attempts at developing platforms focused on full digital preservation started with the development of a reference model for an OAIS and subsequent adoption as standard in 2002, (Flathers, 2017). Stanford university also developed LOCKSS: Lots of Copies Keep Stuff Safe on the premise that the key to keeping a file safe is to have lots of copies. LOCKSS uses a peer-to-peer network for sharing copies of digital material (Maniatis, et al, 2005). Trehub, et al, (2019) describes LOCKSS as a community based network that allows a network of libraries to keep digital materials safe for one another in exchange for access. It was initially designed for preserving e-journals, but today it is used for preserving web content around the world (Maniatis, et al, 2005). It ensures fixity of content through multiple LOCKSS boxes with overlapping content that periodically carry out checksums to compare commonly held content. If the content of one box is different from the others, it is assumed to be damaged and automatically replaced with a good copy. LOCKSS is a cost effective system suitable for bit preservation of web pages but

is hampered by limited types of materials it can preserve as well as the lack of active preservation policies.

In 2003, the Florida Center for Library Automation (FCLA) started the development of DAITSS (Dark Archives in the Sunshine State) software for the Florida Digital Archive (FDA) as an attempt to improve preservation methods (Caplan & Guenther, 2005). The software consists of a set of services that include ingest, data management, archival storage, and but lacks a user interface (Caplan, 2008). The DAITSS focused on preserving master copies with the aim of providing access upon request (hence the dark archives). The FDA ensures long term render ability of digital objects by selecting popular file formats and developing a preservation plan for each of these formats that describes their migration strategy (Caplan, 2008). Format normalization, bit level preservation and forward file migration are the major strategies utilized. DAITSS was released to the public in 2010, but both the repositories and website are offline as from December 2020 due to the decommissioning of DAITSS and the FDA by FLCA in December 2018 (Kemenade, 2020).

In 2010, the National Library of France developed a repository called the Scalable Preservation and Archiving Repository (SPAR), which had the capacity to store 1.5 petabytes of data, (Fauduet & Peyrard, 2010). SPAR was designed to be a full OAIS repository meeting all digital preservation needs (Delaney & de Jong, 2015). SPAR only ensures only bit level preservation. Kemenade, (2020) opines that the major strength of SPAR lies in metadata management where a full description of all digital objects ingested is given. The utilization of metadata standards in the information packages, such as METS, PREMIS, Metadata for Images in XML Standard (MIX) or text MD (an XML

schema that details technical metadata for text-based digital objects) ensures that sense can be made of the diverse types of documents in the repository (Fauduet & Peyrard, 2010). SPAR also introduced the concept of tracks. A track is a collocation of digital objects with the same preservation requirements (Fauduet & Peyrard, 2010) based on characteristics such as legal, political as well as technical requirements. Each track consists of multiple channels. A channel is a collection of objects require similar treatment (Kemenade, 2020).

The concept of digital preserving as a service (DPaaS) has emerged as a solution to some of the digital preservation challenges. DPaaS has been described as a scalable, local or cloud-based third-party service that can be provided at different levels to meet an institution's storage needs, security demands, access requirements, and degree of preservation desired (McGuirk, 2016). DPaaS provides long-term digital stewardship of files by offering access to the digital objects independent of the hardware and software that originally created them in order to “relieve the records owners of the onus of engineering and provisioning the preservation infrastructure” (Nguyen & Lake, 2011). One of these solutions is archivematica that was developed in 2010 as an open source solution for digital archiving with a goal to reducing the cost and technical complexities involved in the establishment of a comprehensive digital preservation solution anchored on best practices (Blewer, et al, 2019). Archivematica provides an integrated suite of open-source tools that enables the processing of digital objects to comply with OAIS functional model (Garderen, et al 2013). The platform supports emulation, as well as migration by monitoring risks to file formats. Normalization is the default preservation



strategy for archiving and is effected at the point of ingest there by converting file formats into preservation formats (Van Garderen, et al 2013).

The choice of a digital preservation program influences the strategies that will be adopted. Although Myntt (ed) (2019) advises digital preservation and archival communities to adopt open technology he also acknowledges that there are instances where proprietary technologies would be more appropriate for information centre with limited IT capabilities. According to Quisbert (2006) long term digital preservation platforms should proactively support forward migration, to ensure that information is compatible with a newer environment. True digital preservation should ensure that digital information resources are readable, usable, findable and trustworthy as long as needed. Any platform that an information center intent on long term digital preservation adopts should support all these aspects or be capable of integrating tools that support them. In addition to a preservation strategy, systems must be able to support integrity checks, versioning, geographical distribution, scalability, preservation metadata standards among others (Kremser, 2012). Most institutional repository software in use are more useful as digital collection management tools but are limited in their support for long-term digital preservation (Ruusalepp & Dobрева, 2012).

## **2.5 Challenges to Digital Preservation**

Libraries, archives and museums today face numerous challenges among them technology obsolescence, information overload as well as maintaining trust as repositories that hold documentary evidence of scholars and citizens (The Council of Canadian Academies, 2015). In addition, digital records are inherently software-dependent posing immense challenges to its long-term preservation (Dar & Ahmad,

2017). In order to make the digital data contained in a bit stream accessible and usable, specialized software and hardware are needed to interpret and process the data according to the file format it is encoded in. This process is known as rendering the bit stream into an understandable information object. Without the specific hardware and software information in digital format is incomprehensible, useless and irrelevant, (Howell, 2000). Technology obsolescence is the single most major risk to long term preservation of digital information, (Hedstrom & Montgomery, 1998) that is accentuated by the technology's continuous flux that gives little timeframe for organisations to migrate digital contents to new software / hardware (Arora, 2009). In addition, Johnstone (2020) notes that there are dozens of carrier formats - floppy disks, hard drives, CDs, DVDs, thumb drives, tapes among others requiring hardware that is no longer manufactured or supported by modern personal computer architectures and software that can no longer be found online, even if the original manufacturer still exists. Today, manufacturers are phasing out optical readers and librarians have no other option but to provide alternatives to access information in CD ROMs. Niehof, et al (2018) and Langley (2019) decried the limited time afforded to managers of digital content to save them due to rapid changes in technology and complexities of technologies involved.

Fragility of the media is another major challenges that custodians of digital material wrestle with'. According to Hedstrom and Montgomery, (1998), the media digital materials are stored on is inherently unstable and without suitable storage conditions and management can deteriorate very quickly even though it may not appear to be damaged. Arora (2006) adds on that most storage devices, without suitable storage conditions and proper

Management, deteriorate very quickly and this may lead to loss of data. Tallman & Work (2018) summed it up by arguing that physical degradation of hard drives and other digital storage media can lead to file corruption.

Rapid changes in technologies render file formats obsolete. This is compounded by the existence of many types of formats making it difficult for repository administrators to keep track of them. Pearson and Webb (2008), described file format obsolescence as a major risk factor threatening the ongoing usefulness of digital information collections. According to Arora (2006), the problem is compounded by the fact that many of the most commonly used computer applications rely on proprietary native file formats to create, save, store, manage and retrieve digital content. According to Duff, et al(2006), digital preservation is an extremely complex, evolving field that requires a great deal of knowledge to understand which when coupled with the high speed of technological changes mean that few organizations are able fully to articulate what their needs are in this area, much less employ or develop staff with appropriate skills. It has also been noted that there is little in the way of appropriate training and "learning by doing" is often the most practical interim measure a point that was brought out clearly by Engelhardt, (2013) in *The DigCurV Review of Training Needs in the Field of Digital Preservation and Curation* done in Germany, Ireland, Italy, Lithuania, and the UK. The study concluded that there was a severe lack of professionals with the skills and competences necessary to deal with digital preservation tasks not only among the existing staff in institutions but also among potential staff in the labour market compounded with the lack of appropriate training options. A recent study by Masenya and Ngulube (2019) in South Africa also identified a lack of training and skills in the industry and went on to recommend use of

external expertise and investment in staff training especially through workshops and seminars. This is in agreement with Alison, et al, (2019) when they shared the University of Glasgow's *Digital Preservation Journey 2017-2019* and found out knowledge and skills among the staff were enhanced by participating in a national digital preservation pilot project and learning from practitioners through workshops and information exchange.

The integrity of digital information resources is critical to their use. Information security encompasses a number of aspect such as: confidentiality, availability, integrity and authentication (Campisi, et al, 2009). In a repository perspective, availability ensures that the information is available and can be used when needed a notion supported by NISO (2007) who described availability as ensuring that the collection is accessible and usable upon demand by an authorized but goes on to clarify that this does not mean that materials should be free and unrestricted for all but attempts should be made to ensure that resources are as widely available as possible within the required constraints. Integrity ensures that it is not corrupted while authenticity ensures that it remains original. Owens (2018) argues that information security goes beyond having backups but also ensuring that any changes done on the digital document is recorded and legitimate.

Intellectual property rights have been described as one of the biggest problems in digital preservation (RLG, 1996). Whitt (2017) opines that unless an information resource is exempted by copyright law; is in the public domain or preservation is done by copyright owner then institutions have challenges because irrespective of the strategy adopted, some form of copying is done leading to infringement of copyright. Kilbride and Norris

(2013) warns that there are risks when digital preservation actions are delayed due to IPR issues and recommends adoption of appropriate and practical actions to manage the risks. Limited funding has been and continues to be an obstacle to digital preservation forcing organisations to develop policies that can be implemented with limited budget (Owens, 2018). Wittenberg et al (2018) noted a trend in higher education funding where limited funds received by the libraries made them focus on content acquisition at the expense of digital preservation activities. Nigerian libraries have failed to invest in infrastructure that support digital preservation due to limited funding (Solomon & Soyen, 2021). Weinraub et al (2018) summed it up by declaring that funding remained a major impediment to the establishment of a robust digital preservation program.

Digital objects today are made of various components all residing in one repository (Wittenberg et al, 2018). This complexity underscores the need to create, capture and maintain more information on context and relationships in order to support discovery, access and rendering in future. Baillieul, et al(2018) notes that versioning which is one of the components of complexity needs authoritative monitoring and proposes infrastructure that supports distributed systems ranging from author personal websites, publisher portals and IRs.

Heterogeneity: heterogeneity refers to the diversity and complexity of digital objects and the systems and technologies used to create, store, and access them. Digital objects come in a wide range of file formats, structures, and encoding schemes, and are created and used on a variety of hardware and software platforms. No single community or organization can to create, collect, and/or preserve just a single type of born-digital or

digitized object since there exists many types of file formats as well as a number of variants of the same format (Johnstone, 2020). Protecting these complex resources combination of technical expertise, organizational planning and policies that organisations cannot provide on their own.

## **2.6 Digital Preservation Trends**

Today organisations implement digital preservation either in a centralized or distributed model. Centralized preservation refers to preservation activities managed by single institution like DAITSS, a digital preservation software application developed by the Florida Center for Library Automation (FCLA) and used by Florida Digital Archive (FDA), a long-term preservation repository service provided by the Florida Virtual Campus for the use of the libraries of the eleven publicly-funded universities in Florida, (FCLA, 2011). The preservation protocol implemented by DAITSS combines bit-level preservation, format normalization, and forward format migration (Caplan, 2008). Decman and Vintar (2013) proposed a centralized digital preservation solution for e-government based on cloud computing. They suggested a single repository for digital information held by governments institutions and departments instead of each managing their own. This way, government institutions would be able to save costs on preservation infrastructure and avoid duplication of efforts.

On the other hand, distributed preservation refers to preservation activities managed by multiple institutions replicating and/ or geographically locating collections. Distributed digital preservation involves collaboration between like-minded institutions and is meant to alleviate immense challenges associated with digital preservation as well as establish communities of practice for adoption of best practices (Skinner & Halbert, 2009;

Weatherburn, 2017)). Skinner and Halbert (2009) proposes that the best practice for distributed preservation is to have: a minimum of three sites for the resources; sites preserving the same content should not be within a 75 to 125 miles radius of one another; sites should be away from paths of natural disasters; should not share power grids; should be under different administrators and should be on live media that is frequently checked for bit rot.

The Florida Centre for library automation (2011) argues that no one institution can preserve everything and neither can they be wholly isolated independent organizations. Collaboration has been hailed as an essential ingredient in the preservation of digital resources (Kalusopa, 2018). According to Dollar & Ashley (2020), digital preservation is a costly and complex undertaking that requires collaborations among all stakeholders terming it a shared responsibility.

There has been a National and international focus on mechanisms for cooperation, coordination, and federation of preservation efforts, as well as the development of shared standards. For example, in the United States, the National Digital Information Infrastructure and Preservation Program (NDIIPP) program is taking the lead in establishing a distributed digital preservation network while in Europe, the PLANETS project funded by the European Union is a major vehicle for integrating distributed preservation services (Baucom, 2019). Proponents of distributed systems Trehub & Halbert, (2012) argue that there is safety in numbers and support the adoption of the Lots of Copies keeps Stuff Safe (software) that supports keeping of many copies.

Corrado and Moulaison (2014) view the preservation of digital information as an economical, managerial, and technological challenge and as such requires decisions on how and where to store. Service models have emerged to provide digital preservation services to publishers and institutions (Wittenberg, et al 2018). This has been attributed to limited expertise as well as failure to keep up with the required skills set required for digital preservation by individual digital repositories, (Cunningham, 2018). In addition, the use of cloud-based facilities increases opportunities for collaboration and offer potential cost-savings over locally hosted solutions in digital preservation, (Pillen & Eckard, 2022).

## **2.9 Evaluation of Digital Preservation Practices**

The main objective of a digital preservation program is to be able to assure stakeholders that information held in digital form and is understood today can be transmitted into an unknown system in the future and still be correctly understood (Antunes et al, 2012). To achieve this objective quality management is one of the essential parts of any digital archive (Dobratz et al, 2010). The idea of establishing the effectiveness of digital preservation practices was born in December 1994 when the Commission on Preservation and Access and the Research Libraries Group created the Task Force on digital archiving with the mandate to investigate the means of ensuring continued access indefinitely into the future of records stored in digital electronic form (Waters & Garrett (eds.), 1996). The taskforce identified a gap that required for a process of certification for digital archives to cultivate a climate of trust about the prospects of preserving digital information (Waters & Garrett (eds.), (1996). From this evolved the first check list which later became ISO 16363/TDR to evaluate digital preservation environments in digital repositories (Baucom,



2019). Lin et al (2020) insists that trustworthiness is not a one-off achievement and should be sustained with regular audit and certification. According to Maemura, et al (2017) assessments should form a major component of a digital preservation program because it is only after an assessment that a digital repository can demonstrate efficacy as well as bridge any gaps identified and consequently promote trust amongst the stakeholders.

Many digital repositories are today using preservation needs analysis to audit the status of their preservation activities. Durant (2019) describes a preservation needs assessment as a broad, holistic evaluation of how an organization cares for and preserves its collections aimed at describing the impact of the existing conditions and policies on the collections, and providing corresponding short, medium, and long-term steps that an organization can take to benefit the materials under its care. Maemura, et al (2017) notes that evaluations can range from simple check lists to rigorous audits while the categories of assessment used can be both descriptive or qualitative (Durant, 2019). Despite the necessity to carry out assessment, it is faced with some challenges with Antunes et al (2012) decriing the lack objective assessable (measurable) features regarding long-term aspects proposing the use of indicators to show the degree of trustworthiness. Researchers like Pearson and Coufal (2013) proposed some metrics by describing an ideal digital preservation environment as consisting of a mix between policies, processes and resources that comprise staff and technologies.

In the area of institutional repositories, Dell and Shultz (2014) felt that it was important that IRs adopted practices that could be tracked, audited and measured as this encouraged the designated communities to trust them as custodians of their research output.

According to Altman et al (2019) much content has been lost due to organizational failure and recommends an audit of content, and the evaluation of organizations themselves to mitigate the risks. Although the preservation community has made considerable progress towards articulating the practices and behaviors of trustworthy preservation, organizations have been slow in conducting reliable evaluation of their digital repositories (Altman, et al, 2019). Maemura, et al (2017) identified three approaches to the evaluation of digital preservation in organisations. These include evaluations aimed at planning, improvement and those aimed at certification. Whatever approach the organisation chose, the evaluation assists the digital repository to document its digital preservation success, report areas that need further growth, and identify challenges that could prevent that growth (Tapscot, 2019).

Although the role of evaluations in digital preservation cannot be faulted, few research in Africa have focused on this. Masenya and Ngulube (2019) focused on need for the adoption of best practices in Nigeria while Umana's (2020) research investigated the long-term digital preservation activities in Namibia. Adjei et al (2019) evaluated digital preservation practices within Ghana while in Kenya Moseti (2016) and Erima et al (2016) focused on identifying the digital preservation practices. Digital preservation is context-specific because it depends on a variety of factors that can differ significantly between organizations, institutions, and even regions. This necessitates for research in the different contexts to indicate the status of long-term preservation in these environments.

## 2.7 Summary

This chapter has reviewed literature related to the efficacy of digital preservation practices. The chapter began with a description of the theoretical frameworks. A review of ISO 14721:2012 and ISO 16363:2012 and their applicability to the study is provided. The empirical literature review starts by reviewing the role of IRs in digital preservation. The discussion picks on differences by scholars on role of institutional repositories (IRs) in digital preservation with some believing that IRs should prioritize improving access, usage, and impact, while others argue that digital preservation is an important function of IRs. The agreement is that IRs should be guided by its stated purpose and the environment in which it operates. The discussion focuses on IRs responsible for storing and maintaining scholarly content for the long term and notes concern that they lack the characteristics of digital preservation repositories and instead focus on access to content.

Further, the chapter looks at digital preservation practices, that include: content characteristics such as type, selection, file formats as well as ingest procedures and their implications to long term preservation. Other practices identified include adoption of digital preservation policies, digital preservation planning; identification of digital preservation strategies, as well as the use of a preservation platform that supports all the above practices. A number of challenges are discussed and since most of these challenges stem from limited finance for development of critical infrastructure, the chapter goes on to expound on the use of collaborations to mitigate these challenges. The extensive survey of literature points to the fact that while there is a growing body of literature on digital preservation practices globally, especially in Europe, Asia and UAS, there is

limited literature on the same within the African context with the chapter closing on a recommendation for more research in this area.

## **CHAPTER THREE**

### **RESEARCH METHODOLOGY**

#### **3.0 Introduction**

Rajasekar, et al (2006) describe methodology as a systematic way to solve a problem. It includes the procedures by which researchers go about their work of describing, explaining and predicting phenomena. Research methodology has also been described as the systematic procedures by which the researcher starts from the initial identification of the problem to its final conclusions and describes its role as ensuring the research is carried out in a scientific and valid manner (Singh, 2006). This chapter examines the research methodology that was utilized to carry out the study and explains the reason for employing the discussed research techniques. The chapter goes further to discuss the data collection tools utilized. An explanation of choice of method is also provided.

#### **3.1 Research Paradigm**

A research paradigm has been defined as “the entire constellation of beliefs, values and techniques shared by members of a given scientific community” (Tombs & Pugsley, 2020, p. 1). This means that these beliefs are shared by a scientific community and guide how communities of researchers’ act with regard to inquiry. Different research paradigms are based on different philosophical foundations and notions of reality. According to Ryan (2018) research views can either be subjective or objective. Objectivity takes the position that there is a single version of what is real, regardless of the researcher’s perspective while subjectivism takes into account the multiple and varied perspectives of what may be real. Research paradigm is the lens that a researcher uses to examine the methodological aspects of their research in order to determine the research methods that

will be used for data collection and analysis (Kivunja & Kuyini, 2017). Four sets of philosophical assumptions constitute a paradigm: axiology, ontology, epistemology and methodology (Mertens & Wilson.,2012). To answer the epistemological aspect of their research, researchers draw from four sources of knowledge: intuitive knowledge (knowledge from beliefs, faith, and intuition); authoritative knowledge (people in the know, books, leaders in organizations); rationalist epistemology or logical knowledge (driven by desire to know the truth); empirical epistemology (wish to understand knowledge through experiences, demonstrable or objective facts), (Kivunja & Kuyini, 2017). Epistemology instills trust in a researcher's data as well as influencing how they go about uncovering knowledge, (Kivunja & Kuyini, 2017). Ontology relates to the nature of existence of reality (McGregor, 2018). Ontology relate to the assumptions made in relation to the kind and nature of reality and what exists (McGregor, 2018; Richards, 2003) that helps a researcher to orientate their thinking about the research problem, its significance, and how they might approach it so as to answer their research question, understand the problem investigated and contribute to its solution (Kivunja & Kuyini, 2017). On the other hand, methodology relates to the research design, procedures and approaches used in the research while axiology are the ethical considerations in the research (Kivunja & Kuyini, 2017).

The four major trends in world views include the positivist research philosophy, interpretivist research philosophy, pragmatist research philosophy, and realistic research philosophy (Zukauskas et al, 2018). The positivist research philosophy claims that the world can be viewed in an objective way (Zukauskas et al, (2018) and that knowledge can be obtained through observation and subsequent verification of the same in

circumstances where the said knowledge is true (Stern, 2004). Positivists differentiate between scientific and normative statements and argue that normative statements cannot be measured by senses and therefore only scientific statements are true (Dawadi, et al, 2021). Positivism research paradigm relies on deductive logic, formulation of hypotheses, testing those hypotheses, and mathematical calculations to derive conclusions. It provides explanations and makes predictions based on measurable outcomes. Positivism has been associated with quantitative methods but in recent times post-positivism has accepted a little bit of subjectivity by utilizing both qualitative and quantitative methods, (Dawadi, et al, 2021).

The interpretivist research philosophy claims the world can be viewed in a subjective manner. Advocates of the interpretivist philosophy argue that it is important to understand the differences among people as opposed to collecting data among objects (Saunders et al, 2007). Da Silva et al (2018) present the interpretivist paradigm as one that seeks a participants understanding of a situation but fails to examine the conditions that lead to certain meanings in some reported experiences nor explain the unintended consequences of an action all significant forces in the construction of social reality. The interpretivism paradigm is aimed at finding new interpretations. Zukauskas et al, (2018) postulated that the philosophy places honors on the researcher since he/she has a specific role to play and goes on to add that the research will be based and depend on the interest of the researcher and thus the subjectivity. Tombs & Pugsley (2020) stress the importance of empathy for the researcher to make interpretations of different meanings individuals and groups attach to the activities and their accounts of the activities.

The realism research philosophy is the belief that reality is shaped over time by values such as social values, political values, cultural values or gender based values, (Ryan, 2018). According to realists, reality resides or is independent of the human mind. It is a world view that takes the position that there is no possibility of attaining a single correct understanding of the world (Maxwell, 2012). Realism is often associated with quantitative methods but it is compatible with qualitative methods, (Madil, 2008).

The Pragmatist paradigm believes in facts (Zukauskas et al 2018). Yin (2010) describes pragmatism as a worldview that supports the selection of appropriate research methods in relation to research questions being studied. Similarly, Saunders et al (2007) adds that the proponents of this philosophy put forward that the research question determines the approach used by the researcher. To a large extent the paradigm is applicable in mixed methods (Creswell & Clark, 2011; Saunders, et al 2007). Never the less, researchers adopting a pragmatist position have the liberty to choose those research methods or strategies that can best answer their research questions (Creswell, 2007). For researchers adopting a pragmatic approach, the most important question is whether the research has helped to find out what the researcher wants to know and guided by this gather all sorts of data in order to best answer the research question, (Dawadi, et al, 2021). Epistemologically, pragmatism is premised on the idea that research can steer clear of metaphysical debates about the nature of truth and reality and focus instead on practical understandings of concrete, real-world issues (Patton, 2005). Pragmatism places emphasis on interrogating the value and meaning of research data through examination of its practical consequences (Morgan, 2014).



Although qualitative research is often associated with the interpretivism paradigm, Goldkuhl (2012) argues that there are alternatives especially in qualitative information systems research where pragmatism is widely adopted. Infact, Mackenzie and Knipe (2006) opine that several paradigmatic positions such as positivist, constructivist, interpretivist, transformative, emancipatory, critical, de-constructivist and pragmatic can be adopted for the conduction of information systems research. According to Morgan (2014), pragmatism can serve as a philosophical paradigm for social research, regardless ofwhether that research uses qualitative, quantitative, or mixed methods. This is a view also held by Yin (2009) who put forward that researchers utilizing this paradigm may choose to use a quantitative method or a qualitative method, or to conduct a mixed methods study using both kinds of methods, all depending on which choice best befits the research questions. Kelly and Cordeiro (2020), identify an emphasis on actionable knowledge, recognition of the interconnectedness between experience, knowing and acting and inquiry as an experiential process as the three major principles that underpin the pragmatic approach. Pragmatists argue that the purpose of an inquiry is basically to create knowledge in the interest of change and improvement and the essence of a pragmatist ontology is actions and change (Goldkuhl, 2012). Pragmatism informs qualitative research in two ways. One is inquiry into practical questions in search of useful and actionable answers and two in the making of pragmatic method decisions based on situations and opportunities during the enquiry process (Patton, 2014). Plath (2013) while supporting the use of pragmatic paradigm in evaluative qualitative research notes that the approach allows researchers to gather evidence from a range of sources and to critically evaluate them in terms of their strengths, limitations, and applicability to the

practice setting. In addition, evaluators must maintain balance, fairness, and neutrality (Patton, (2014) lending the pragmatic paradigm useful for evaluators who would not wish to be caught in between subjectivity and objectivity. At the same time, as a qualitative inquiry framework, pragmatism directs researchers to seek practical and useful answers that can solve, or at least provide direction in addressing concrete problems (Patton, 2014). Godfrey-Smith, (2002) summed this up by positing that knowledge should bring about a positive change in organisations. Knowledge should benefit an organisation by enhancing the discovery of truth about policies, practices and decisions (Cavaleri, 2008; Visser, 2019). A pragmatist approach, with its epistemological focus on the inquiry, process and practicality, was found to be more useful for this study that sought to holistically evaluate the appropriateness and effectiveness of digital preservation practices in support of long term care of digital resources. According to Rescher (2000) pragmatic research is applicable to research that cares not just for the efficiency of means but for their appropriateness as it involves combining a whole range of evaluative factors including but not limited to efficiency and effectiveness. The study was able to achieve this by ensuring proximity of the phenomenon as well as the triangulation of sources and methods that pragmatism allows (Goldkuhl, 2004). The choice of pragmatism as a philosophy for this study was strongly influenced by:

Pragmatism emphasizes the practical aspects of research and evaluation. Its focuses on finding solutions that work in real-world contexts and generating practical knowledge that can be applied to improve systems or processes. In establishing the efficacy of digital preservation strategies, it allowed the researcher to assess the digital preservation

practice's effectiveness, usability, and impact on users' experiences, thereby providing actionable insights for any improvement on digital preservation.

Secondly, pragmatism acknowledges the importance of multiple perspectives and diverse viewpoints. In qualitative evaluation, understanding the experiences, perceptions, and behaviors of users and stakeholders is essential. The paradigm enabled the researcher to consider different perspectives and explore the practical consequences of those perspectives on the research process. It allowed for the inclusion of diverse voices, ensuring a comprehensive understanding of the digital preservation's programs strengths, weaknesses, and potential improvements.

Pragmatism recognizes the importance of collaboration and engagement with stakeholders. In order to establish the efficacy of digital preservation practices, the researcher was required to interact with key stakeholders. Pragmatism encourages researchers to actively engage these stakeholders throughout the research process, incorporating their perspectives, needs, and concerns into the assessment. This collaborative approach enabled the researcher to enhance the validity and relevance of the research findings and in addition enhanced the likelihood of the ownership and acceptance of potential changes and recommendations by the stakeholders.

### **3.2 Research Approach**

Creswell (2014, p. 31) defines a research approach as plans and the procedures for research that span the steps from broad assumptions to detailed methods of data collection, analysis, and interpretation". Singh (2006) describes a research approach as choice of an investigator about the components of his project and development of certain

components of the design. From these two definitions it can be deduced that a research approach is like a guide to the researcher on the best way possible to achieve his/her objectives.

There are three types of approaches; qualitative, quantitative and mixed methods. Creswell (2009) defines quantitative research as a means for testing objective theories by examining the relationship among variables. These variables, in turn, can be measured, typically on instrument so that numbered data can be analyzed using statistical procedures. Mixed methods research utilizes both the qualitative and quantitative research approaches. It is anchored on the premise that the integration of qualitative and quantitative data yields additional insight beyond the information provided by either the quantitative or qualitative data alone, (Creswel & Creswel, 2018).

Qualitative research is a means for exploring and understanding the meaning individuals or groups ascribe to a social or human problem (Creswell, 2009). The process of research involves emerging questions and procedures, data typically collected in the participant's setting, data analysis inductively building from particulars to general themes, and the researcher making interpretations of the meaning of the data. Qualitative research is anchored on understanding the meaning and context of the phenomenon under study. According to Patton (2014), human beings engage in efforts meant to make the world better. Information systems are attempts at making the world better and their success or failure is influenced by the organizational context (Kaplan & Maxwell, 2005). Viray (2015) defines Information System (IS) as a set of interrelated components working together to collect, retrieve, process, store, and disseminate information. Evaluation of these systems help human beings to understand whether their objectives

have been met. Hennink, et al, (2020) describe information systems as a combination of information technology, procedures and people and advise that studies should be centered around these three components with the primary purpose of discovering new insights that can aid the effective application of information's systems in organisations. Qualitative approach is useful when an evaluator wishes to study issues that are not easily partitioned into discrete entities, or to examine the dynamics of a process rather than its static characteristics (Kaplan & Maxwell, 2005).

Qualitative data in an evaluation provides information and generates findings that are useful to people about processes and outcomes for decision making (Patton, 2014). Patton goes on to add that a researcher wishing to adopt the qualitative approach in evaluation seeks to answer questions such as: what constitutes program quality? what is excellence? what is the quality of the program's design, implementation and results with the aims of either increasing quality and excellence, enhancing effectiveness or generating and demonstrating a model of quality

The qualitative research approach was deemed suitable to the study establishing the efficacy of digital preservation practices in institutional repositories for several reasons. Qualitative approach allows for an in-depth exploration of complex phenomena. Digital preservation practices in institutional repositories involve various factors such as organizational policies, technological infrastructure, and user engagement. A qualitative approach enabled the researcher to delve into these factors, understanding the underlying processes, challenges, and experiences related to digital preservation. Through methods like interviews, observations, and document analysis, the researchers was able to gather rich and detailed data that provide insights into the effectiveness of these practices.

In addition, qualitative research is well-suited for understanding the perspectives and experiences of individuals involved in digital preservation. This research topic involves multiple stakeholders, including repository administrators, librarians as well as IT support staff. By employing qualitative methods, the researcher was able to the diverse viewpoints of these stakeholders and gain a comprehensive understanding of their beliefs, attitudes, and behaviors regarding digital preservation. This holistic perspective was crucial for assessing the efficacy of the practices from different angles and identifying potential gaps or areas for improvement.

Furthermore, qualitative research allows for flexibility and adaptability in exploring emergent themes and issues. The field of digital preservation is dynamic and constantly evolving, with new challenges and technologies emerging over time. A qualitative approach enabled the researcher to respond to these changes, adapting their research questions and methods to capture the latest developments and insights. It also allowed for the exploration of unexpected findings or novel practices that may not have been initially anticipated, contributing to a deeper understanding of digital preservation efficacy.

Qualitative research enabled the researcher to uncover contextual factors and complexities associated with digital preservation practices. Institutional repositories operate within unique organizational, cultural, and technological contexts. By conducting qualitative research, the researcher was able to investigate how these contextual factors influence the implementation and effectiveness of digital preservation practices. This understanding was crucial for the development of targeted recommendations and strategies that are tailored to the specific institutional environment.

Qualitative research facilitated the exploration of the human and social dimensions of digital preservation practices allowing the researcher to examine not only the technical aspects but also the social, cultural, and organizational aspects that impact the efficacy of these practices. By capturing narratives, motivations, and interactions, the research shed light on the human factors involved in digital preservation, such as collaboration, knowledge sharing, and decision-making processes.

### **3.3 Research Design**

According to Yin et al (2018) all types of empirical research study has research design. Bhatacherjee (2012) describes a research design as a comprehensive plan for data collection in an empirical research that acts as a blueprint to enable the researcher to answer specific research questions or testing specific hypotheses. A design's purpose is to avoid situations where evidence or data collected does not address the research questions, (Yin, et al, 2018). Creswell and Creswell (2018) identified five types of qualitative research designs. These are: grounded theory, phenomenology, narrative research, ethnographies and case study research designs. Leedy, et al (2019) describes grounded theory as a design where the researcher focuses on a process related to a particular topic including people's actions and interactions—with the ultimate goal of developing a theory about the process. Phenomenology design emphasizes the study of conscious experiences to assist the researcher understand the reality surrounding the phenomena by utilizing human judgment, perceptions, and actions of the participants (Bhatacherjee, 2012). The narrative design is rooted in the disciplines of psychology, sociology and anthropology and focuses on the recollections and stories of individuals who have had experiences related to these phenomena (Leedy et al, 2019). The

ethnography design enables a researcher to study an intact cultural group in a natural setting over a prolonged period of time by utilizing the interview and observation data collection methods (Creswel & Creswell, 2018). Dul and Hak (2008) describes a case study as a study in which (a) one case (single case study) or a small number of cases in their real life context are selected, and (b) scores obtained from these cases are analyzed in a qualitative manner.

### **3.3.1 Case Study**

This study sought to establish the efficaciousness of digital preservation practices in Institutional Repositories (IRs) of selected public universities in Kenya in order to propose best practices for digital preservation in IRs. The study utilized a multiple case study design incorporating the evaluation research aspect. Weis (1998) describes an evaluative research as a systematic assessment of the operation & or the outcomes of a program or policy, compared to a set of explicit or implicit standards & means of contributing to the improvement of the program or policy, while Powell (2006, p102) defines it as “a type of study that uses standard social research methods for evaluative purposes, as a specific research methodology, and as an assessment process that employs special techniques unique to the evaluation of social programs”. Christie and Alkin (2013) identified three branches of evaluation: Methods, Use and Values. Mertens and Wilson (2014) added the fourth branch of Social Justice. Evaluative research involves immediate stakeholders as part of the evaluation process from the beginning. Powell further explains that a researcher carrying out evaluative research can utilize standards which are guidelines or recommended practices developed by a group of experts that serve as models for good service. There are methodologies developed for evaluative



research with Stern (2004) categorizing these methodologies into: the criteria or standards based position, which is concerned with judging success and performance by the application of standards; the causal inference position, which is concerned with explaining programme impacts and success and the formative or change oriented position, which seeks to bring about improvements both for programmes and for those who participate in them. This study is more of the first category basing its methodology on a criterion (ISO 14721:2012 and ISO 16363/TRD:2011) to establish the effectiveness of digital preservation practices within selected university libraries in Kenya.

Connaway and Powell (2010) stress that a case study is appropriate for investigating phenomena when (1) a large variety of factors and relationships are included, (2) no basic laws exist to determine which factors and relationships are important and (3) when the factors and relationships can be directly observed. At the same time Yin (2018) asserts that case studies are applicable in evaluative studies in four ways. These are: to explain the presumed causal links in real-life interventions that are too complex for the survey or experimental strategies, to describe an intervention and the real-life context in which it occurred, case studies can illustrate certain topics within an evaluation in a descriptive mode and to enlighten those situations in which the intervention being evaluated has no clear, single set of outcomes. He goes on to say that the case study can be used to answer the “how” question of a research problem. Case studies have been known to involve a lot of interrelationships none of which is deemed more important than the others. Case studies are suited to provide an understanding of the interactions between information technology (IT)-related innovations and organizational contexts thus it was deemed a suitable research design as the success of digital preservation is dependent on the

organizational context (Williamson & Johanson, 2017). The use of more than one case allowed for a wider view of digital preservation within the case context and also to enlighten and answer the question “how” effective the digital preservation practices adopting Yin (2009) recommendation for such questions. Institutional repositories are information technology innovations whose study cannot be separated from the context of the organisation. Yin and Campbell, (2018) while exalting the strengths of multiple case studies argued that research utilizing this design was considered of higher quality than single cases providing a more robust data that can be relied upon.

For this study, multiple case studies allowed for a comprehensive examination of the research topic. Digital preservation practices vary across different institutional repositories, influenced by factors such as organizational culture, resources, and policies. By conducting multiple case studies, the researcher was able to capture this variation and gain a holistic understanding of the efficacy of digital preservation practices. This enabled the exploration of commonalities and differences among cases, identifying patterns, challenges, and success factors that may not be apparent in a single case study.

In addition, multiple case studies provided a basis for comparison and cross-validation of findings. By examining multiple institutional repositories, the researcher was able to compare and contrast their digital preservation practices, outcomes, and effectiveness. This comparative analysis allowed for the identification of best practices, lessons learned, and areas for improvement. The triangulation of data from multiple cases strengthens the validity and reliability of the research findings, as consistent patterns or themes emerging across cases enhance the confidence in the conclusions drawn.

Lastly, multiple case studies enabled the researcher to increase the reliability of the findings since the multiple cases support the transferability of findings to other contexts. The insights gained from studying multiple cases can be more easily generalized to other institutional repositories and inform the development of guidelines, recommendations, and best practices for digital preservation. By examining the efficacy of digital preservation practices in different settings, the researcher was able to identify common principles and strategies that could be applied in various institutional contexts, contributing to the advancement of the field as a whole.

### **3.4 Population**

The population or universe has been defined by Singh (2006) as the entire mass of observations, which is the parent group from which a sample is to be formed. Sekaran (2003) describes it as the entire group of people, events, or things of interest that the researcher wishes to investigate. He goes on to identify population element as a single member of the population where as a population frame is the listing of all the elements in the population from which the sample is drawn. This study involves “establishing the efficacy of digital preservation practices in institutional repositories (IRs) of selected public universities in Kenya”. The population for this study includes all public universities in Kenya as indicated Open DOAR (2020) listing. According to Open DOAR only 22 public universities had been listed and these formed the target population.

### **3.5 Target Population**

Moffat (2015) defines target population as all the members in the group to whom the investigation is related. Asiamah et al (2017), point out that the general population contains some elements that lack the desired characteristics for the study. They go on to

advice researchers undertaking qualitative research to focus on participants who can best share experiences and thoughts in order to address the qualitative research goal by identifying and eliminating individuals of the general population who may not have the ability to share experiences and thoughts in ample clarity and depth. According to Asiamah et al (2017) not general population can participate in the study. Public universities who have an institutional repositories and have registered their policies and mandates with the Registry of Open Access Repository Mandates and Policies (ROARMAP) formed part of the target population since the policy and mandate were an important aspect of the study. Only 11 out of the 22 have been able to do this (ROARMAP, 2020). A summary of the population type, characterizes and size is given in table 5 below.

**Table 4: Population Type Characteristics and Size**

<b>Type of population</b>	<b>Criteria for inclusion</b>	<b>Population size</b>
General population	Public universities with established institutional repositories as listed in OpenDOAR	22
Target population	Public Universities who have registered their open access policies and mandates with ROARMAP	11

### **3.7 Sample Design and Sampling Techniques**

Onwuegbuzie and Collins (2007) define sampling designs as representing the framework within which the sampling occurs, including the number and types of sampling schemes and the sample size. A good sample design should therefore result to a truly representative sample; must result to a small sampling error; must be viable in the context

of funds available for the research study and should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence (Taherdoost, 2016).

Taherdoost (2016) further identifies two types of sampling techniques: probability sampling and non-probability sampling. Probability sampling is based on the concept of random selection, whereas non-probability sampling is 'non-random' sampling. According to Saunders et al (2007), non-probability sampling is suitable for case studies. Expounding on this, Gerring (2007) explains that the goals of case study research dictate that they are met through purposive (nonrandom) selection procedures and that a case is chosen based on the way it is placed within the population. For this reason, then, non-probability sampling techniques were adopted. Criterion purposive sampling was used in selecting to select three public universities. Palinkas et al (2015) argued that purposive sampling is mostly used for the identification and selection of information-rich cases related to the phenomenon of interest. Purposive sampling allows a researcher to identify and select respondents who are knowledgeable and have experience with the phenomenon under study. More so it is applied in qualitative research to allow for the in-depth study of the phenomenon. The three cases were selected purposively using the following criteria: length of time they had been in existence; how far they had been able to develop their institutional repository infrastructure judged by a number of factors such as: Time the repository was registered with OpenDOAR (UoN; 2013, KU; 2012, JKUAT; 2013); amount of content as indicated in their repository websites at the time (over 3500 items) and registration with ROARMAP. The criteria on time of registration was based on the assumption that the universities had well-established IRs with content that may

necessitate the need for active digital preservation while registration with OpenROAR indicated evidence of policies which was one of the elements under study.

The library and ICT departments were also chosen purposively owing to the fact that the library department is responsible for the management of the institutional repository and that the ICT department is involved in one way or another in the management of the technical aspects of the IRs.

Expert sampling a type of purposive sampling was used in selecting the actual participants from the library and ICT departments. From the library, these included repository administrators, senior library management comprising of university librarians or director and the deputy university librarians responsible for policy development and implementation selected from the library. From ICT department, the technical person attached to the library was selected.

### **3.8 Sample Size**

Sekaran (2003) defines a sample as subset of the population and comprises some members selected from it. According to Guest, Namey and Chen (2020), data saturation is the most commonly employed concept for estimating sample sizes in qualitative research. Braun and Clarke (2021) have described data saturation as the point at which both the quality (richness, depth, diversity and complexity) and quantity are achieved. Qualitative research utilizes a small sample in order to support the depth of case-oriented analysis, (Asiamah, N. et al, 2017). In addition, qualitative samples are purposive, that is, selected by virtue of their capacity to provide richly-textured information, relevant to the phenomenon under investigation (Vasileiou, et al 2018). In this study, in-depth interviews

were conducted allowing the researcher to gather significant amounts of data from a small sample both in terms of cases included in the university and the actual respondents.

Data coding was guided by the interview and document analysis guides as well as the observation guide allowing the researcher to collect enough in-depth data for all the themes. The researcher went over and over the data to identify any content that could have been left out otherwise. In the UoN one Library director, two deputy university librarians, two repository administrators, a systems librarian and one ICT library liaison person formed part of the sample. For university JKUAT and KU the sample consisted of one university librarian, one deputy librarian (each had one), two repository administrators, the systems librarian and one ICT library liaison person in each case.

A total of 19 (nineteen) respondents were involved in this study, seven from UoN and six from each of the other universities under study as shown in table 6.

**Table 5: Sample Population Framework**

	University Librarian	Deputy University Librarian	Systems Librarian	IR staff	ICT staff	Total
UoN	1	2	1	2	1	7
JKUAT	1	1	1	2	1	6
KU	1	1	1	2	1	6
Total						19

### 3.9 Data Collection Techniques/Methods

Data collection is a process of collecting information from all the relevant sources to find answers to the research problem (Kabir, 2016). The main methods of qualitative data

collection are participant observation, qualitative interview, focus groups and qualitative observation (Bryman, 2012). Zikmund et al (2012) describe a focus group interview as an unstructured, free-flowing interview with a small group of people usually between six and ten that allows for multiple perspectives. While this is one of the major strengths of focus groups, Berg and Lune (2017) cautions against the use of focus groups to collect data on events, behaviors, or feelings rather they recommend the use of one-on-one interviews.

The interview method was used to gather information from all the respondents. A Qualitative interview is as a conversation whose purpose is to gather information from an interview by an interviewer (Berg & Lune, 2017).

### **3.9.1 Face to Face Interviews**

Osang et.al. (2013) have described face to face interviews as a series of questions a researcher addresses personally to respondents. Patton (2002), describes three types of qualitative face to face interviews. These are: the informal conversational interviews; the general interview guide approach and the standardized open ended interview. An interview guide is a list of questions to be explored and is prepared to ensure the same line of questions is taken with all interviews. The aim of this research was to evaluate the digital preservation practices in the selected universities and the interview guide enabled the researcher to be able to focus on the topic under study thus collect quality and comprehensive data. Cohen et.al. (2007) recognized the importance of an interview guide by outlining its benefit as being able to increase the comprehensiveness of the data as well as helping the researcher to anticipate logical gaps in data and close them in advance.



Face to face interviews can be structured where questions are clearly defined or unstructured where questions are influenced by the responses the researcher receives. To be able to achieve this, the perspectives of the participants was necessary.

Unstructured interviews have the advantage of allowing flexibility, enabling the researcher to adjust the questions depending on responses given in order to collect adequate data (Walliman, 2011). This is because the interviewer was able to ask probing questions and also the respondent had a chance to ask for clarifications in case of misunderstanding. Interviews made it possible to obtain the data that was required to meet specific objectives of the study. This was achieved by enabling the researcher to be up close to the sources of data as well as support the triangulation of sources. Probing questions enabled the researcher to ensure that data relevant to the research questions was collected. The face to face interview also enabled the researcher to clarify a few questions since some of the interviewees were not familiar with some of the digital preservation aspects enabling the respondent give the relevant response.

Interviews give the interviewer a chance to explain the purpose of the research thereby convincing the respondents of its importance. This way, the respondents were able to cooperate fully in the research process.

According to Qu and Dumay (2011), Many respondents do not like to expose their negative side, an interview allows an interviewer to get such information through interaction and genuine conversation and continued reassurance of confidentiality thus allowing the researcher to get more insights into the phenomenon under study.

### **3.9.2 Observation**

Zikmund et al (2013) is a systematic process of recording behavioral patterns of people, objects, and occurrences as they happen as well watching and recording events as they happen. Cohen (2007) describes observation as a powerful tool for gaining insight into situations although he does not recommend its application on its own due to validity and reliability.

He recounts several strengths of the method including: its ability to eliminate bias if accurately applied, Secondly, the currency of information obtained, that is information obtained under this method relates to what is currently happening; its simplicity devoid of complications emanating from past behavior or future intentions or attitudes, and lastly its independence of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents. All these qualities were found to be desirable for the study. Observation can be participatory or non-participatory. Non-participant observation is where the researcher watches either openly or concealed but does not engage in the activities while participant observation is where the researcher is also a member of the group being observed, (Morgan, et al, 2017).

This study utilized non-participatory observation that enabled the research to observe actions directly. Observation went hand in hand with the face to face interviews and was used to confirm responses from these interviews. It was also performed together with document analysis to identify whether what policy proposes corresponds with what the database has on file format, type of content and metadata.

Non participatory observation was used to collect data in the following areas: Types file of formats, metadata creation, quality control activities, digital preservation strategies, types of contents, digital preservation platform, adherence to policies and procedures and digital storage media. An observation guide was used with themes derived from the research objectives (Appendix 2).

### **3.9.3 Document Analysis**

Gross (2018) defines document analysis a form of qualitative research that uses a systematic procedure to analyze documentary evidence and answer specific research questions. And that document analysis can be used as a standalone or as a component of qualitative research to triangulate data collected through other methods. Documents can either provide primary or secondary data. Documents providing primary data would include minutes, policies emails and photographs among others. Bowen (2009) sums up the purpose of document analysis as gathering background information and determining if implementation of the program reflects program plans.

In this study, document analysis was used to collect primary data in order to support triangulation. Institutional repository policies, procedures and the website were analyzed to collaborate data gathered through the interviews. Gross (2018) postulates that when document analysis is used in triangulation, the documents corroborate or refute, elucidate, or expand on findings across other data sources. This helps to guard against bias. In this study, document analysis was used in answering questions on institutional responsibility and viability especially in regard to policies, plans, budgets and mandates. At the same time, the repository database and website was analyzed to confirm responses given about

content, metadata and file formats. A document analysis guide (appendix 2) was used in collecting the required data.

### **3.10 Reliability and Validity**

Reliability refers to the ability of research data to be consistent, dependable or replicable. Bashir, et al (2008 citing Joppe, 2000) define reliability as: “The extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable” (p. 1). Validity refers to the correctness or precision of a research reading. Validity is concerned with whether the findings are really about what they appear to be about (Saunders et al 2007).

Bell et al (2018) citing LeCompte & Goetz (1982) decried the challenge of achieving validity in qualitative research as it mostly relies on case studies and small samples. According to Golafshani (2003 citing Stenbacka, 2001) the concept of reliability is even misleading in qualitative research and proposes that reliability be replaced with truthfulness. Equivalent psychometric measures of data gathering tools used in qualitative research in place of reliability and validity include triangulation, transferability, credibility, trustworthiness, truth, value, applicability, consistency and confirmability.

#### ***Trustworthiness***

Trustworthiness points out to the level of confidence readers have in what a researcher has reported. Research procedures utilized by a researcher during the research activity and in reporting assists in creating trust. Qualitative researchers are encouraged to adopt strategies such as triangulation, transferability, dependability, reflexivity and

conformability in order to ensure the trustworthiness their research. (Bell, et al, 2018). The strategies adopted for trustworthiness in this study are discussed below.

### ***Triangulation***

Triangulation refers to using more than one particular approach when doing research in order to get richer, fuller data and/or to help confirm the results of the research (Wilson, 2014). In the context of qualitative research, Leedy et al (2019) describe triangulation as the collection and comparison of multiple kinds of data, with the goal of finding consistencies or inconsistencies among them. It is an attempt to gain more than one perspective on phenomenon that is being investigated. In fact, Yin & Campbell, (2018) caution against the reliance of one source when utilizing a case study design. It may be two or more perspectives, despite the triangulation. The researcher used triangulation of instruments where for example respondents were asked questions and then observation and document analysis checked whether the information given tallied with what was on the ground.

### ***Transferability***

Transferability has been described as “Extent to which a research study’s findings might be similar or applicable to other individuals, settings, and contexts” (Leedy et al, 2019). Qualitative research findings lean towards contextual uniqueness due to the small sample size utilized and raises the question on whether the results can hold the same in another context (Bell, et al 2018). This study ensured transferability by describing the accounts under study in detail. This is in line with Bell et al (2018) who citing Geertz (1973) advised qualitative researchers to give thick descriptions of the phenomena under study in

order to provide others with a database for making judgements about the possible transferability of findings to other contexts.

### ***Dependability***

Leedy et al (2019) describe dependability as a concept that accounts for the ever-changing contexts within which research studies take place and thus requires researchers to provide in-depth descriptions of their data collection methods. Bell et al (2018) proposed both a detailed description of the research methods and a peer audit throughout the research. This research sought to achieve dependability by giving a detailed description of the research methods used as well as frequently consulting the supervisors to act as auditors.

### ***Reflexivity***

According to Leedy et al (2019) reflexivity is the ability to actively identify personal, social, political, or philosophical biases that are likely to affect the ability of a researcher to collect and interpret data and then take whatever steps they can to reduce such influences. An important dimension of this capacity is to step away from what they have thus identified, while acknowledging their own roles in the conduct of their research. This research sought to achieve this by jotting down the comments of both the researcher as well as the respondents, reading through the interpretations and editing out evidence of bias, memoing immediately after an interview as well as meticulously describing all choices of research methods.

### ***Confirmability***

Leedy et al (2019) have described confirmability as the ability of the researcher to make a concerted effort to base their conclusions on their actual data as much as possible. The

researcher was able to achieve confirmability by using audio recordings to complement what was written in note book as well as taking screenshots from the database. This provided an audit trail that enabled the researcher to look back and confirm whether the findings discussed corresponded with the data from the field.

### **3.11 Data Analysis**

Cohen, et al, (2007) describes qualitative data analysis as the process of organizing, accounting for and explaining data and involves making sense of the participant's definition of situation taking note of patterns themes and categories. Qualitative data analysis focuses on text rather than numbers and are therefore not straight forward to analyze Bell, et al (2018). According to Bryman (2012), qualitative data can be analyzed using techniques such as grounded theory, thematic analysis, Content analysis, discourse analysis and narrative analysis.

Content analysis has been described by White and Marsh (2006) as a highly flexible research method that has been widely used in library and information science. Content analysis involves summarizing and reporting written data (Cohen et al (2007). There are two types of content analysis: conceptual analysis and relational analysis (Wilson, 2016). Conceptual analysis involves coding content to certain words, themes or themes? Whole. Relational analysis involves identifying the relationships between the words and themes (Wilson, 2016). Concept analysis is used to systematically analyze documents obtained or generated in the course of research. Content analysis is based on the premise that many words from interviews, observations and documents can be reduced to or organized into categories in which words or word units (paragraphs) share the same meaning. One of the strengths of qualitative content analysis is that it allows the researcher to scrutinize data

closely, identifying concepts and patterns. Through this process, the researcher may come across some patterns and concepts that were not foreshadowed but are important to consider. It has been argued that in such cases, it is acceptable for the researcher to alter his/her interests and research questions to pursue these new patterns (Leedy et al, 2019). Hsieh and Shannon (2005) identified three approaches to content analysis: conventional, directed, or summative. Conventional content analysis involves deriving coding categories directly from the text data while directed analysis involves starting with a theory or relevant research findings as guidance for initial codes. A summative content analysis involves counting and comparisons, usually of keywords or content, followed by the interpretation of the underlying context (Hsieh & Shannon, 2005). Conventional content analysis is also known as inductive content analysis and is used when the research purpose is to describe a phenomenon on which there is only limited literature and no existing theory (Zaidman- Zait, 2014). On the other hand, directed content analysis is aimed at validation or extension of a theoretical framework (Mayring, 2014). Researchers using the directed content analysis approach identifies the device that will guide them through data collection and analysis (Kibiswa, 2019). The researcher is guided by categories and themes that they had defined prior to the data collection during data analysis and interpretation. Hsieh and Shannon (2005) recommend directed content analysis when the researcher wants to extend a theory in a context or a situation different from the one in which that theory was developed, to provide supporting or non-supporting evidence for an existing theory (e.g., to validate the theory), to complete the description of a phenomenon already theorized or to use the existing theory to guide the discussion on the research findings.



This study sought to establish the efficacy of digital preservation practices of selected IRs in public universities in Kenya. The study utilized directed content analysis where existing research theoretical frameworks relevant to the study informed the codes and themes used for analysis. The choice of directed content analysis was informed by the fact that the study was evaluative in nature and required an evaluation criterion that was adopted from ISO 14721:2012 and ISO 16363: 2011. Predefined themes identified from documents such as policies, websites, interview responses and observation included: content characteristics, policy, planning, digital preservation strategies were borrowed from the standards. Challenges of digital preservation and recommendations were new themes that the researcher identified from face to face interviews.

The researcher utilized Kibiswa, (2019) three phase of directed qualitative content analysis for data analysis as described below.

*Phase 1: Preparation Phase*

*Step 1: Developing the study's frame and operational definitions*

This step involved studying the ISO 14721:2012 and the ISO 16363:2011 and identifying the themes and sub themes as shown in table 6 below:

**Table 6: Themes and Subthemes as Derived from Theoretical Framework**

Standard	Theme	Subtheme
ISO 14721:2012	Content	Type of content
		Acquisition and selection
		File format
		Metadata
ISO16363:2011	Policy	Evidence of policy
	Plan	Evidence of plan
	Preservation strategies	
	Policy	Evidence
		Content
	Plan	Responsibility for preservation
		Evidence of plan
		Technology
		Finance
	Preservation strategies	Staffing
Procedural documentation		

*Step 2: Determining the unit of analysis and sampling materials to be analyzed*

In this phase the researcher analyzed data from policies, procedures, websites, databases, notes face to face interviews, transcribed interviews as well as notes from observation. The researcher also settled on the unit of analysis. A unit of analysis has been defined as “whole text which is large enough to be considered as a whole and small enough to be kept in mind as a context for the meaning unit during the analysis process” (Graneheim & Lundman, 2004). The theme was chosen as the unit of analysis.

*Step 3: Getting a sense of the data*

This involves searching and locating the themes and sub themes identified in step 1. The researcher read through the source materials highlighting the words, phrases or other text

that are identical, close or similar to the themes and subthemes identified in step 1. By going through the data back and forth until data saturation is achieved. See table 8.

***Phase Two: Data Analysis Phase***

*Step 4: Data coding and organizing*

This involved reading the material interpretatively in order to assign codes to the highlighted text indicating the source in order to give context to the data. See table 8

*Step 5: Making connections, interpreting them, and drawing conclusions*

The researcher analyzed the data and presented it in a meaningful way, offering supporting evidence through a table summarizing the responses and results obtained from content analysis. Visual aids such as screenshots, photographs, and verbal citations were utilized. Additionally, two new themes emerged from the data analysis, focusing on the challenges of digital preservation and potential solutions to address these challenges.

*Step 6: Verifying interpretations plausibility and ensuring trustworthiness.*

The researcher effectively communicated the research findings, accompanied by a clear articulation of the research methodology, including the rationale for its selection. The interpretations and conclusions drawn were supported by quotations, figures, and tables, providing evidence for the reader to assess the transferability, dependability, and confirmability of the findings. This approach empowers the reader to form their own judgments regarding the research outcomes.

***Phase Three: Reporting the Analysis Process***

*Step 7: Making an appropriate outline for a detailed presentation.*

The structure of the thesis is guided by the school of post graduate thesis guidelines. The findings for the study are discussed guided by the themes identified during data analysis.

*Step 8: Thick description of the research history and findings.*

The researcher has provided a detailed description of the findings by presenting the findings, interpreting the findings within the context of their origin, placed the findings within the context of existing literature, made conclusions based on the findings as well provided recommendations guided by the findings.

**Table 7: Sample Directed Content Analysis**

Research Question	Data from policy interview and observation	Code	Theme as identified from ISO 14721 and ISO 16363	subtheme1	subtheme	subtheme 3
1.What are the characteristics of digital objects accepted for storage in IRs in the public universities under study	<p>Content Policy: The repository shall accept submissions of the following types of materials: -----</p> <p>Any research which includes a confidential report for a sponsor, i.e., company/commercial third party, will not be included, or will have restricted access unless otherwise agreed by the sponsor.</p> <p>Authors who are depositors shall only submit their own work for archiving. Where an item has multiple authors: At least one author must be a member of staff or a student of , and The submitting author shall obtain the permission of the co-authors</p> <p>“we take content as guided by our OA policy but sometimes because of PC targets we accept outside the define types. For example, we had started ingesting <b>photographs of management</b> but we found this unsustainable”</p>	<p>content</p> <p>policy</p> <p>Procedures</p> <p>Defined Content</p>	<p>define the type of data to be collected</p> <p>should have supporting policies and procedures that govern the selection criteria</p> <p>set procedures for negotiating for the information</p> <p>define content</p>	<p>Type of content</p> <p>Type of content (deviation noted)</p>	<p>Acquisition and selection</p> <p>Defined content boundaries</p>	<p>Procedures for submission</p>

### **3.12 Ethical issues**

Hammersley and Traianou (2012) describe ethics as set of principles that embody or exemplify what is good or right or allow us to identify what is bad or wrong. Wiles (2013) classify ethical frameworks as consequentialist, principlist, ethics of care and virtue ethics and expound that consequential ethics are based on consequences of actions, principlist are non-consequential and relate to principles to respect participants' autonomy such as informed consent, ethics of care are based on compassion and the need to act to the best of interest of the people involved in research while virtue ethics focus on the moral character of the researcher.

Based on these frameworks then, this study took into account several ethical issues such as:

#### **Confidentiality**

Confidentiality is concerned with maintaining privacy and respect for autonomy and sets to ensure that information given to another person will not be repeated without their permission (wiles, 2013). Hammersley (2012) recognized the importance of confidentiality in research as it is meant to protect the participants from negative consequences if the information is revealed. The researcher ensured confidentiality and privacy by ensuring that the respondents names did not appear anywhere in the published research. Any information collected from one participant was not repeated to another participant. During document analysis any information that was not for the public was treated as such.

**Plagiarism**

“Plagiarism involves either (a) presenting another person’s work as being one’s own or (b) insufficiently acknowledging and identifying the sources from which one has drawn while writing” (Leedy et al, 2019, 102). Bellet al (2018) advise researchers to avoid plagiarism at all cost and this was taken into consideration throughout this research by ensuring that any information borrowed from other researchers was adequately cited and referenced according to APA 6th edition referencing style. More so the work was checked for any unintentional plagiarism by using Turnitin plagiarism checker before submission.

**Voluntary and Informed Consent**

According to Hammersley and Traianou (2012), the principle of informed consent is rooted in the idea that individuals should have the freedom to make decisions based on what they believe is best for themselves. This principle serves as the foundation for obtaining informed consent from participants in research studies. As advised by Yin (2018), researchers should ensure that participants are fully informed about the nature of the study and actively seek their voluntary participation.

Informed consent plays a crucial role in ensuring that individuals have a clear understanding of what it entails to participate in a specific research study, enabling them to make a conscious and deliberate decision regarding their involvement. Following Yin's (2018) recommendation, the researcher in this study explained the purpose of the research to the participants and ensured that they were aware of the time commitment and the type of information required during the interviews. By providing this information, the

researcher aimed to obtain informed consent from the participants, acknowledging their autonomy and ensuring they were well-informed before deciding to take part in the study.

### **Adherence to Legal and Policy Regulations**

In accordance with the research regulations outlined by Wiles (2013), researchers are obligated to comply with legal requirements pertinent to their studies. In line with this, the researcher in this study adhered to the necessary procedures and obtained a research license from the National Commission for Science, Technology, and Innovation (NACOSTI), as indicated in the appendix. Additionally, authorization was sought from the universities under investigation, in accordance with their individual research policies. The researcher obtained letters of approval from these institutions, which are included in the appendix, demonstrating compliance with their regulations and commitment to ethical research conduct.

### 3.13 Summary

This chapter provides an overview of the research methodology employed in the study. The research paradigm and approach were discussed, with the study being guided by the pragmatic paradigm, which is particularly suitable for qualitative research seeking practical answers to research questions. To enhance the quality of the data collected, multiple case studies were utilized.

The population for this study consisted of public universities that had established Institutional Repositories (IRs), and the target population specifically included public universities with IRs that had open access policies. Purposive sampling was employed to select three universities as well as the participants within those universities who would be involved in the study.



Data collection in this qualitative research study involved interviews, observation, and document analysis. Directed content analysis was employed to analyze the qualitative data. To ensure the credibility of the research findings, several strategies were employed, including triangulation, flexibility in data collection, confirmability of the analysis, transferability of the results, and dependability of the research process.

Ethical considerations were carefully addressed throughout the study. Measures were taken to maintain the confidentiality of participants, avoid plagiarism, and adhere to legal and policy requirements. These ethical considerations were vital to protect the rights and well-being of the participants and maintain the integrity of the research process.

The chosen methodology facilitated the collection of relevant data aligned with the research objectives and research questions. Table 7 provides a summary of how the research methodology aligned with the specific research objectives and questions, demonstrating the appropriateness and effectiveness of the chosen approach.

**Table 8: Linkage of Research Objectives and Questions to Sources of Data**

Objective	Research Question	Source of data
Determine the characteristics of digital resources in the institutional repositories	What types of digital objects are accepted for storage in IRs in the public universities under study?	Librarians and IR administrators, ICT Liaison and Database and website
Audit the existing digital preservation policies	Are the digital preservation policies suitable to support long term digital preservation?	Librarian, IR administrators, ICT Liaison, document review
Appraise the digital preservation plans of the selected IRs	How adequate is the digital preservation plan?	Librarian, IR administrators, ICT Liaison, document review
Examine the digital preservation strategies practiced by the IRs in the selected universities	To what extent do the preservation strategies being used by the IRs adequate to support long term preservation?	Librarian, IR administrators, ICT Liaison, document review, Database
Identify digital preservation challenges in the selected public universities	What challenges does the IRs experience in regard to preserving research outputs?	Librarian, IR administrators, ICT Liaison.
Make recommendations and propose best practices for digital preservation in The IRs of the selected universities.	What can be done to improve the effectiveness of digital preservation practices in the selected universities ?	Librarian, IR administrators, ICT Liaison, best practices from existing literature

## CHAPTER FOUR

### DATA PRESENTATION, ANALYSIS AND INTERPRETATION

#### 4.0 Introduction

This chapter analyses, presents and interprets the data. The aim of the research was to evaluate digital preservation practices in Institutional Repositories (IRs) of selected public universities in Kenya with the aim of making recommendations as well as propose best practices for digital preservation in IRs. Data was collected from face to face interviews, analysis of policies and procedures and form observation. The study generated primary qualitative data that was analyzed manually using directed/ deductive content analysis. The process identified key themes and sub themes guided by ISO 14721:2012 and ISO 13663: 2011 as proposed by Hsieh & Shannon (2005). The analysis resulted into the following key themes: Characteristics of information materials accepted into the IRs, Preservation policy, Preservation plan, digital preservation strategies in place, the challenges experienced and proposed solutions to digital preservation. Data was presented in narrative form and tables where applicable.

#### 4.1 Response Rate

This study interviewed 19 people involved in the establishment and day to day management of institutional repositories with a 100% response rate. The respondents were distributed as shown in table 8 below.

**Table 9: Respondent Rate by Participant Categories**

University Libraries	PARTICIPANTS' CATEGORY					Total
	A	1	2	1	2	
B	1	1	1	2	1	6
C	1	1	1	2	1	6
Total						19

#### 4.2 Characteristics of Materials Accepted into the Repositories

The research sought to find out the characteristics of resources accepted in the institutional repositories of the universities under study. Themes identified in this objective included type of content, file formats, procedures for ingest as well as metadata.

A summary of the responses discussed in this section are indicated in table 11 below:

**Table 10: Summary of Responses on Characteristics of Digital Materials Accepted in the IRS**

Question	A	B	C
Are the boundaries for the digital collection clearly defined in an acquisition policy?	Yes (following types of materials shall be accepted)	Yes(repository shall accept submissions of the following types of materials	Yes (repository shall accept submissions of the following types of materials
What formats of information resources are accepted in the repository?	PDF (including image files) VOB (video)	PDF, MP4 (video)	Defined in policy (table 14). In addition mp4 and vts, html
Are these formats clearly articulated in a policy?	No (File formats that are platform-independent, vendor-independent,	No mention of file formats	File formats clearly defined

	non-proprietary, stable, widely supported		
Are their guidelines for depositors and ingest?	No guidelines but in process of developing. Policy on submission only identifies who can deposit	No guidelines. Policy on submission only identifies who can deposit	No guidelines. Policy on submission only identifies who can deposit
What type of metadata is collected during ingest?	Descriptive metadata (from policy) and database Technical (file type and size) Rights metadata (permissions to access digital objects, use of creative commons-policy clear and license used Administrative (permissions to manage digital objects, persistent identifiers, metadata documenting changes like corrections to the original)	Descriptive metadata (from policy) and database Technical (file type and size) Rights metadata (permissions to access digital objects, no open access license identified, Administrative (permissions to manage digital objects, persistent identifiers)	Descriptive metadata (from policy) and database Technical (file type and size) Rights metadata (permissions to access digital objects, no open access license identified, Administrative (permissions to manage digital objects, persistent identifiers)
Is the metadata collected enough for long term preservation?	Lacked technical metadata such as software and hardware descriptions, File type and size automatically identified by system	Lacked technical metadata such as software and hardware descriptions, File type and size automatically identified by system	Lacked technical metadata such as software and hardware descriptions, File type and size automatically

			identified by system
What metadata standards do you apply?	Dublin Core Metadata	Dublin Core Metadata	Dublin Core Metadata

#### 4.2.1 Types of Digital Materials

The study found that some of the resources were common in all three universities under study. These resources included books, conference/workshop/seminar papers, theses and dissertations, policies/ reports/ newsletters, public lectures and speeches, and journal articles, graduation resources (video and lists). There were a few types of resources that were not common to all of them such as multimedia and undergraduate projects as well as a collection called archives that had microform thesis with only metadata, pictorial collection and rare collections.

The study established that in all the universities under study, a large part of the content was taken by thesis and dissertations distributed shown in table 12 below:

**Table 11: Distribution of Content Type across the University**

University	A	B	C
Thesis and dissertation	43557	10951	2800
research / journal articles	31758	5527	259
conference proceedings	7308	598	1515
Books/book chapters	2049	266	
Others	26037	584	188
Total	110709	17926	4762

At university A, thesis and dissertations formed 39.4% of the total content, while, research papers/journal articles were 28.7%, Conference proceedings (6.6 %) books and book(1.85%), others (23.5). In B, thesis and dissertations formed 61.1 %, journal articles (30.8%), conference proceedings (3.3) %, books and book chapters (1.5%) and others(3.3%). The distribution of content type in university C was as follows: thesis and dissertations formed (58.8%), research papers (5.4%), conference proceedings were (31.8%) and others (3.9%). From this then, it can be deduced that 78% of the content in the IRs were primary information sources (thesis and dissertation, conference proceedings and journal /research articles) that are very useful in scholarly communication for furthering research. The large concentration of thesis, dissertations and research papers in all the repositories underscored the need for the IRs to consider their preservation to ensure that future generations have access to these resources.

The findings established that thesis and dissertations had an analog equivalent mainly because it was mandatory that postgraduate students presented both a soft copy and a hard copy before graduation. Still a good number of the items in the collections had only the digital copy especially journal articles published in open access journals as well as journal articles published in journals that the library had not subscribed to. It was evident from the responses that the availability of an analog equivalent impacted on digital preservation negatively posing risks to born digital content such as research articles published in e-journals. Respondents considered the existence of an analog copy was security enough against loss of information since the IR could always digitize in case of loss.

#### 4.2.2 Defined Boundaries for Materials accepted

The study established that there were clear boundaries defined on which resources were accepted. This was presented as a section in the open access policies documenting the type of content that was accepted for storage by the repositories. These resources are as shown in table 13 below:

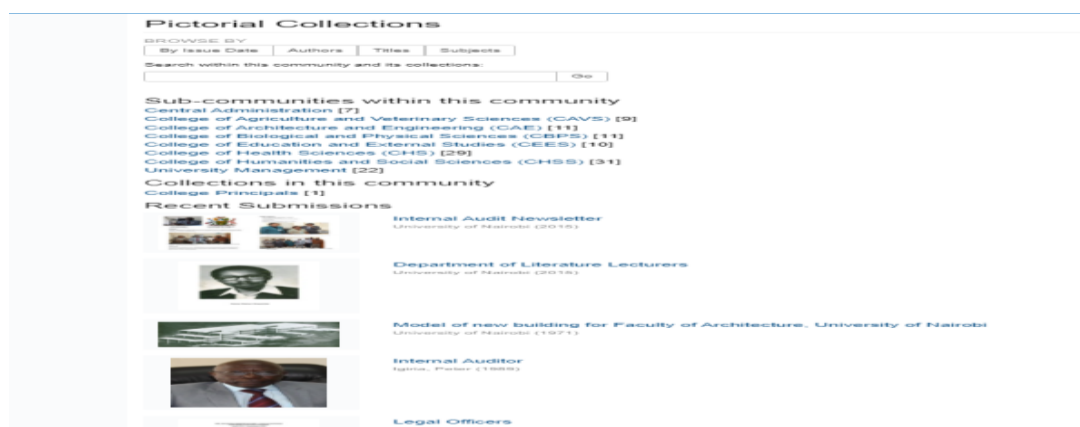
**Table 12: Type of Content Received in the IRs of the Selected Universities**

TYPES OF MATERIALS	UNIVERSITY		
	A	B	C
Theses and dissertations	✓	✓	✓
research projects	✓	✓	✓
Scholarly Research articles (published peer reviewed and pre-prints	✓	✓	✓
Open lectures	✓	X	X
Conference/workshop Proceedings;	✓	✓	✓
Books and book chapters	✓	✓	X
Monographs	✓	X	X
Image collections (paintings, pictures, drawings, illustrations	✓	X	X
Audio and audio-visual materials	✓	X	X
Technical reports and working papers	✓	✓	X
Inaugural lectures, distinguished lectures, speeches	✓	X	X
Datasets	✓	X	X
Refereed designs	✓	X	✓
Creative, performance-based and visual arts outputs that have research components	X	X	✓
Admissions lists	✓	X	X
Graduation lists	✓	X	X
University policies	✓	X	X
Events programs	✓	X	X
Valedictory presentations	✓	X	X
University calendars	✓	X	X
University magazines	✓	X	X
Forms	✓	X	X
Newsletters	✓	✓	X
Literary publications	✓	X	X
Journalism student media content	✓	X	X
Other materials produced by academic/research staff and approved by Deputy Vice Chancellor, (Research, Production and Extension), Directors of, Schools/Deans of	✓	✓	
Learning objects (past papers, lecture notes and presentations)	X	✓	X
Unpublished scholarly works	X	X	✓



Although the acquisition section of the OA policies clearly identified content accepted a clause indicating “any other material” gave leeway to receive all types of content procedurally. This left them open to receive content not well defined such as charters and graduation booklets and undergraduate projects. The implications of the “any other material” was reflected in one repository where a pictorial collection had been established consisting of photographs of university management and other dignitaries scanned and saved in PDF. It was also noted that this collection was deemed unsustainable because the repository management found it difficult to keep up with frequent changes in management. Although new materials could not be added to this collection, no decision had been made on what to do with what had already been deposited in the repository except to make it inaccessible to the public. These instances revealed the importance of a strong selection policy to determine what was ingested into the repository with a goal to long term preservation as well as adhering to the characteristics of a repository of cumulative and perpetual.

Figure 3 below shows an example of a discontinued file collection



**Figure 3: Sample Pictorial Collection**

The findings revealed the need for a needs analysis as defined by the OAIS reference model. The needs analysis enables IR administrators to make suitable decisions in regard to the “other material” as prescribed by the policy by ensuring that they are anchored on user requirements.

It was also noted that the acquisition policy had well defined procedures for withdraw of items from the repositories as indicated through statements such as: “items will be preserved indefinitely”; “items may only be removed from there repository due to: proven copyright violation or plagiarism; legal requirements and proven violations; national security; falsified research; and request by author”. In line with the IR characteristics of cumulative and perpetual, the policies also stated that “withdrawn items will not be deleted but will be removed from public view” and that “withdrawn items' identifiers/URLs will be retained indefinitely”. These statements emphasized the need for strong quality control procedures to avoid cases where the resources were ingested into the repository but are withdrawn for one reason or the other.

#### **4.2.3 Types of Formats Accepted in the Institutional Repositories**

On what type of formats were accepted in the repository, all the universities accepted the Portable Document Format (PDF) for text, JPEG for images, MP3 for audio and MP4 for videos as the preferred file formats. A look at the file formats in the collection revealed that it was true the PDF file format was the most dominant at 97%. This was a clear indication that a large chunk of the resources in the repositories were textual. The video, audio and image formats formed the other 3%. Some video file formats were indicated as unknown and no documentation was available to support their use against formats

documented in OA policy. One repository did not accept video files at all. Lacking documentation to support decision making, no reasons were given for this.

URI: <http://hdl.handle.net/123456789/4574>

Date: 2018-06-07

[Show full item record](#)

#### Files in this item

---



Name: Output\_Merge (2).mp4  
Size: 503.1Mb  
Format: Unknown

[View/Open](#)

#### This item appears in the following Collection(s)

---

- [Graduation Video](#) [6]

URI: <http://hdl.handle.net/123456789/4879>

Date: 2019-01-30

[Show full item record](#)

#### Files in this item

---



Name: VTS\_01\_0.IFO  
Size: 70Kb  
Format: Unknown

[View/Open](#)

#### This item appears in the following Collection(s)

---

- [Graduation Video](#) [6]

#### ***Figure 4: Sample Unidentified File Format***

Figure 4 above shows an examples of unknown file formats in one repository. In addition to the repository software failing to recognize the file formats, the files had different file formats which were not among the preferred formats defined in the policy. These findings reveal a major lapse in digital preservation in that although there was an attempt to limit the number of file formats through policy, the repository did not adhere to the policy and accepted a variety of file formats.

Despite having identified the file formats preferred by the repositories, the policies did not specify distinctions within the same file format families and after investigation some files were found to have different PDF versions such as 1.3 (acrobat .4x) and PDF 1.4 (acrobat .5x).

Respondents exhibited an inadequate depth of knowledge relating to the choice of file formats suitable to support for long-term preservation and access as indicated by the following sample statement.

*“I don’t think file format matters at all. After all, I have been able to open and use my files without any problems for the many years I have used my computer”. RA-B*

The above statement was an indication that even those in charge of the repository did not have basic knowledge of the relationship between file format selection and the repositories could not advice on the same nor normalize during ingest. It was also noted that pictorial objects though created in digital format were converted to PDF format leading to loss of formatting that is an indication of to poor quality as well as exposing the image to integrity queries.

Although all the respondents said that they had clearly defined boundaries for formats, the documentation revealed some gaps. Only one IR had well defined file formats in its policy as detailed in table 14.

**Table 13: File Formats Accepted in the IRs**

<ul style="list-style-type: none"> <li>• Adobe® PDF (.pdf)</li> <li>• Audio and video file formats (. aiff, .aif,.aifc, .tiff, .jpeg, .gif)</li> <li>• Microsoft Office Excel® (.xls)</li> <li>• Machine-Readable Catalogue Records - MARC</li> <li>• Microsoft Office Powerpoint® (.ppt)</li> <li>• Microsoft Office Word® (.doc, .docx)</li> <li>• Moving Picture Experts Group (.mpeg, .mpg)</li> <li>• Text file Formats (HTML, TXT (text), DAT (data: ASCII data), RTF (rich text format), and XML]</li> </ul>
---

Although the table above clearly defined file formats, the policy was not clear on whether the file formats were to be proprietary or nonproprietary.

Another IR policy described a broad criterion for format selection leaving it open for any file format that met the criteria to be used. The statement in the policy read as follows:

*“File formats that are platform-independent, vendor-independent, non-proprietary, stable, widely supported are recommended; The Repository did not accept executable binary files if alternatives are available and Digital Repository staff may convert to more appropriate formats any content that is in obscure or little-used formats for compatibility reasons”. A- OAP, 2012*

Although, the above policy statement lacks specificity in file formats, it gave evidence that the IR recognized the importance of non-proprietary file formats in digital preservation for long-term access. The remaining IR policy was silent on file formats. These findings indicate that the IRs had yet to understand that digital resources had to be managed in a life cycle manner because decisions made at creation such as the choice of the wrong file format could affect its availability and usability in future.

The study sought to find out the suitability of the repository software to manage file formats. It was established that DSpace enabled the repository to designate three levels of preservation format; supported, known and unsupported. It also supports bit preservation by ensuring that the file remains the same through frequent check sums. However, none of the IRs had performed a checksum action nor scheduled for the same.

#### **4.2.4 Procedures for Receiving (Ingesting) Materials in the IR**

All the repositories under study had some guidelines on how content was to be received in the repository but these guidelines only covered textual information. Audio visual and multimedia content was not covered yet the repositories received them. The content was presented to the repository administrators who after checking for quality uploaded the items. Although all the respondents agreed that they had procedures on how the digital content was received and processed, there was no documented evidence and only one IR indicated that it was in the process of preparing a guide for self- archiving. These results establish that IRs were not able to have consistency and quality management strategies that were supported by documentation that is to be used to induct new staff as guided by ISO 16363 / TDR (2011).The implications of poor procedure documentation during ingest was noticed in one IR where staff changes to the section was done frequently and file formats accepted varied as new staff took over duties.

The table below provides a summary of interview responses on procedures for ingest.

**Table 14: Guidelines for Ingest as Reported Verbatim**

University	Type of resource	Guidelines for ingest
A	Thesis	A well labeled CD with both a word and PDF copy of the work. Receive documents via email from departments
	Journal articles	Should indicate whether it is wholly open access or it is still on embargo. The person depositing should be the owner of the intellectual content
	Others	Ownership File format Content
B	Thesis	Thesis checked for quality by school of postgraduate in regard to format and content  Receives well labeled CD from the school of post graduate
	Journal articles	Checked by the heads of department and forwarded to library for uploading
	Others	Received and uploaded not much done to them
C	Thesis	Thesis checked for quality by school of postgraduate in regard to format and content  Receives well labeled CD from the school of post graduate to check referencing and plagiarism
	Journal articles	Checked by the heads of department and forwarded to library for uploading
	Others	Received and uploaded not much done to them

All the three universities received PHD and masters' theses in CD ROM for ingest into the repository. In cases where the final CD ROM was picked from the school of post graduate studies, there were incidences where the repository sometimes received blank CD ROMs. The problems were attributed to weak quality control procedures as well as poor coordination between the source departments and the IR. Storage challenges for the CD ROMs were experienced inform of lack of space as well as storage equipment. In all the IRs, CD ROMs were stored in boxes exposing them to many environmental factors of deterioration thus reducing their life making them unsuitable as backups for the future. Figure 5 below is a sample of the storage conditions for CD-ROMs.



*Figure 5: Sample CDRom Storage Environment*



Some IRs requested for both a PDF and a word document. This decision was informed by the fact that whereas the PDF is the best accepted file format, the word document allowed repository staff to make corrections to the document when needed.

#### **4.2.5 Metadata**

The study found that all the IRs utilized the DSpace repository platform. the study sought to find out the adequacy of the repository software in the manage file formats, the respondents agreed that DSpace was enough to guide them on what types of formats to include since it allowed them to choose three levels of preservation formats; supported, known and unsupported.

This study also sought to find out the metadata schema utilized and the dublin core metadata schema was found to be common to all the IRs. This schema allowed the IRs to create metadata with the following elements: provenance (author), title, date of publication, language, type of resource (thesis, speech.), and the universal resource identifier(URI).

In cases where the resource was a chapter in a book or a journal article in a journal, structural metadata was not provided. Structural metadata ensures that the resource is linked to other related components and prevents separation over time a necessity for long term access. It also allows users to navigate from one section of the complete digital object to the other.

Although there are open access licenses that control the use of open access resources, only one repository utilized the creative common licenses to support its open access publishing method. The license was included with the dissemination metadata in order to

inform users of what is expected of him/her when using the information material in order to avoid cases of infringement.

Technical metadata collected was minimal and included file format and file size. Due to lack of a metadata policy, metadata describing the technical environment was missing. Technical metadata enables IRs to interpret the technical environment to support emulation and migration of digital resources to up to date computing environments.

Administrative metadata collected included rights metadata that indicated whether the item was available on full text to everyone or just a section of the users. One IR was ahead of others and provided an administrative metadata policy indicating that any actions performed on a digital resource had to be documented. This was given in form of a note to show changes that have been made to the original document such as corrections to title, among others (errata and corrigenda lists). Administrative metadata documenting actions on a digital object provides an audit trail to protect the integrity of the object.

From these findings, it can be concluded that in terms of metadata a lot needs to be done to support digital preservation since currently the IRs are concentrating on descriptive metadata to promote access forgetting that future access is dependent on how well the digital resources are preserved.

## Edit Item

Item Status   Item Bitstreams   **Item Metadata**   View Item   Curate

### Add new metadata

**Name:** atmire.cua.enabled

**Value:**

**Language:**

PLEASE NOTE: These changes are not validated in any way. You are responsible for entering the data in the correct format. If you are not sure what the format is, please do NOT make changes.

### Metadata

Remove	Name	Value	Language
<input type="checkbox"/>	dc.contributor.author	AWSC, Women's Economic Empowerment Hub <input type="text" value="acfd594-5311-4e84-833c-"/> <input type="button" value="Lookup"/>	<input type="text"/>
<input type="checkbox"/>	dc.contributor.author	Kimani, Lydia Wanjiru <input type="text" value="9d37#e2-aa35-4373-9485-1"/> <input type="button" value="Lookup"/>	<input type="text"/>
<input type="checkbox"/>	dc.date.accessioned	2022-12-08T12:34:15Z <input type="text"/>	<input type="text"/>
<input type="checkbox"/>	dc.date.available	2022-12-08T12:34:15Z <input type="text"/>	<input type="text"/>

*Figure 6: Sample Metadata Collected via DSpace*

<input type="checkbox"/>	dc.date.issued	2017	
<input type="checkbox"/>	dc.description	Short Video Documentary	en_US
<input type="checkbox"/>	dc.description.abstract	This 34minutes 11seconds-long video was recorded in June 2017. Lyula Kimani was born and educated in Embu. Her father was a District Officer and later a Provincial Commissioner while	en_US
<input type="checkbox"/>	dc.description.provenance	Submitted by dspace admin (periodicals@uoi.ac.ke) on 2022-12-08T12:34:15Z. No. of bitstreams: 1	en
<input type="checkbox"/>	dc.description.provenance	Made available in DSpace on 2022-12-08T12:34:15Z (GMT). No. of bitstreams: 1 license_rdf: 811 bytes, checksum: 6868ccc48a14c8d591352bb9eaf7f8239 (MD5)	en
<input type="checkbox"/>	dc.description.sponsorship	Bill & Melinda Gates Foundation (BMGF)	en_US
<input type="checkbox"/>	dc.identifier.uri	https://youtu.be/nHCZjc1FCBg	
<input type="checkbox"/>	dc.identifier.uri	http://erepository.uoi.ac.ke/handle/11295/162005	
<input type="checkbox"/>	dc.language.iso	en	en_US
<input type="checkbox"/>	dc.publisher	UoN, AWSC, Women's Economic Empowerment Hub	en_US
<input type="checkbox"/>	dc.rights	Attribution-NonCommercial-NoDerivs 3.0 United States	*
<input type="checkbox"/>	dc.rights.uri	http://creativecommons.org/licenses/by-nc-nd/3.0/us/	*

**Figure 7: Sample Metadata Template**

Respondents felt that they collected enough metadata to support long-term digital because for most of the resources, they had had an analog equivalent that could always be used to create another copy if worse came to the worst and therefore they did not have to bother about quality metadata for preservation.

These findings revealed a lack of awareness on the role of metadata in digital preservation may be because librarians are used to descriptive metadata during cataloguing and handling of online resources was new to them. The respondents were not aware of any preservation metadata standards as depicted in the following statements:

*“I feel the metadata collected is enough to enable us manage the resources for however long we wish” RA- A*

*“Dublin core is the best metadata standard for this kind of activity. Although the resources are in digital format, what we do is no different from cataloguing physical books” RA B*

### **4.3 Preservation Policy**

The research sought to find out whether the universities under study had a digital preservation policy and if the policy supported digital preservation. A summary of responses for this theme is provided in table 16 below.

**Table 15: Summary of Digital Preservation Policy Responses**

	Question	A	B	C
1	Does your repository have a digital preservation policy?	✓ (will retain items indefinitely, and ensure continued readability and accessibility and take appropriate measures as warranted; Does not indicate how this will be achieved	✓ (Will endeavour to provide continued readability and accessibility of <b>all</b> items deposited in the repository). Does not indicate how this will be achieved	✓ (will endeavour to provide continued readability and accessibility of all items deposited in the Digital Repository). Does not provide how this will be achieved
2	If yes, does the repository adhere to it?	Yes. deviations noted: Content not defined in the policy was ingested into the repository as well as file formats not defined in policy; Pressure to increase items due to webometrics	Yes: Deviations noted: Content not defined in the policy was ingested into the repository as well as file formats not defined in policy; Pressure for items due to webometrics	Yes: Deviations noted Content not defined in the policy was ingested into the repository as well as file formats not defined in policy
3	Is the intention of long-term preservation well-articulated in the policy?	✓ (provide long-term preservation;	✓ (supports the long-term preservation,	✓ (repository will provide a permanent record of the intellectual output of C, items will be accepted and retained indefinitely, C supports the longterm preservation)
4	Has the policy been revised and if yes by whom?	Has not been revised	Has not been revised	Has not been revised
5	Is the policy clear about the review cycle?	No specific timeline given	5 years	Policy reviewed from time to time, but not later than five (5) years
6	Who is responsible for the policy?	University Open Access Committee reporting to Senate and the vice-chancellor	Office of the vice-chancellor	Office of the DVC Academic
7	Does the policy address rights management issues in relation to digital preservation?	Yes: depositors to sign a depositor agreement	Yes: depositors required to sign a depositor agreement	Yes: depositors required to sign a depositor agreement
8	If no to 1 above, do you plan to have a digital preservation policy?	N/A	N/A	N/A
9	If yes, what is your timeline?	N/A	N/A	N/A

### 4.3.1 Existence of Policy

All the respondents acknowledged the existence of a digital preservation policy.

The study found that the universities had articulated their commitment to long-term preservation of the digital resources as evidenced by the following language within the digital preservation policy

*“The Digital Repository shall be the means for the long-term archiving, preservation and retrieval of materials deposited within it and otherwise provides a permanent record of the University’s scholarly activity by employing the latest technology to aid that objective.” A- OAP*

*“C Digital Repository will: Provide free, searchable access to this output and make possible its long-term archiving and preservation as well as provide a permanent record of the intellectual output of C.” C- OAP*

*“Items Will Be Retained Within B Digital Repository Indefinitely and that B University Will Endeavour to Provide Continued Readability and Accessibility of All Items Deposited in the Repository.” B- OAP*

The section was found to be very basic since the statement to committing the IRs to long term preservation summed up the extent of the digital preservation section. This indicated a lack of a comprehensive digital preservation policy.

### 4.3.2 Adherence to Policy

A look at how well the existing policy was adhered to produced varied answers. The repositories felt that performing backups every now and then was all the policy required of them while few respondents felt that they had partially adhered to policy since the repository accepted materials not indicated in the acquisition policy. From these findings then, it was clear that the IRs were not doing very well on policy and that major gaps existed that weakened their digital preservation efforts.

### **4.3.3 Revision of Policy**

On how often the policies are revised, all the three universities under study indicated that they had not yet revised them. Only one case had indicated the adoption date in their policy. According to the directory of open access, C adopted their policy in April 2012 while B adopted their policy in January 2014.

University A's policy does not give a time frame within which it will be reviewed. Instead it states that "review will be done from time to time", nothing binding. When asked why this is the case the respondents said that since the concept of the IR was "still new" when the policies were being developed, they did not want to commit themselves since they were not sure of the success. This is despite having been in existence for over five years. B and C had given a time frame of "not more than five years". For C this is already overdue and there were no plans in place to review it. Implications of this weakness is based on the dynamic nature of technology that keeps changing. it is important to consider frequent policy review in order to accommodate any emerging issues arising from technology watch.

### **4.3.4 Responsibility for Preservation**

Responsibility for the digital preservation policy was articulated in a general clause within the open access policy with all three cases having open access committees that directly reported to the senate and the Vice-Chancellor. The policy also gave responsibility for the day to day administration of the repository to a repository administrator who reported to the university librarian. Since the digital preservation



policy is only a small section of the respective OA policies of the universities concerned, they lack a section defining the roles and responsibilities in regard to digital preservation.

#### **4.3.5 Rights Management**

The study sought to find out how rights management was carried out and its effectiveness to support digital preservation. The study established that all IRs had ensured that depositors gave rights to the repository staff so that they can undertake long term preservation actions without seeking permissions from them. Depositors were required to sign deposit agreements granting repository staff permission to store, copy and format/manipulate the materials in order to ensure that they can be preserved and made available in the future. Specifically, they granted them the right to: without changing the content; translate the submission to any medium or format for the purpose of preservation and keep more than one copy of this submission for purposes of security, back-up and preservation (sample deposit agreement provided in the appendix). All the respondents indicated a wish to develop a comprehensive digital preservation policy but they could not give a definite time line since it was something they had not thought about before but acknowledged that this research had given them food for thought.

#### **4.4 Digital Preservation Plan**

The research sought to find out whether the selected public universities had a digital preservation plan in place formalizing the preservation actions to be undertaken in support of their commitment for long-term preservation. A summary of responses observations and findings from document analysis for this research objective is given in table 17 below.

**Table 16: Summary of Responses on Adequacy of Plan to Digital Preservation**

		A	B	C
1.	Does the repository have a digital preservation plan?	No	No	No
2.	If yes, does it spell out the responsibilities for preservation?	N/A	N/A	N/A
3.	Is the preservation done in-house or the services are outsourced	In-house	In-house	In-house
4.	Does the plan provide for financial sustainability?	There is no specific plan for the IR. Budgeting done overall for the library	There is no specific plan for the IR. Budgeting done overall for the library	There is no specific budget plan for the IR. Budgeting done overall for the library OA policy indicates 0.8 million for Upgrades and migration, Digitization and preservation processes as well as staff training and skills development but no allocation
5.	Does the budget have a vote dedicated to digital preservation?	No	No	No
6.	Does the plan provide for technology sustainability and technology monitoring?	Although there is no plan, technology monitoring is done though no documentation to support this	Although there is no plan, technology monitoring is done though no documentation to support this	Although there is no plan, technology monitoring is done though no documentation to support this
7.	Does the plan provide for skills development?	No plan but staff attend seminars though not on digital preservation	No plan but staff attend seminars though not on digital preservation	No plan but staff attend seminars though not on digital preservation
8.	Are there mechanisms in place to change the preservation plan, as user needs and technology needs change?	N/A since there is no plan	N/A since there is no plan	N/A since there is no plan

#### **4.4.1 Existence of Digital Preservation Plan**

There was a major challenging in evaluating this aspect as all the IRs under study lacked digital preservation plans. However, some aspects of planning were established through various channels such as budgets and minutes of meetings. Respondents were of the opinion that whatever the plan should entail was already provided for by the open access plans as indicated by the following statements;

*“the depositors’ agreement allows us to copy or migrate resources when need be “A-RA*

*“we have already committed ourselves to preserve the resources” C-UL*

All the same, respondents acknowledged the need for the development of a clear roadmap on how digital preservation was to be undertaken in order to ensure consistency and sustainability.

Lacking a documented digital preservation plan, all preservation actions undertaken were done in-house. The respondents argued that the repositories were still establishing themselves and that with time they may consider more complex preservation actions that may require external services.

#### **4.4.2 Financial Planning**

In regard to finance, there was no budget set aside as such for digital preservation. If finances were all included together in the library budget. This is a great challenge for long-term preservation since it is a resource intensive activity that requires reliable finances for personnel, media and other technologies. The budgets were also diminishing forcing the libraries to focus on access rather than preservation. the study established that due to limited finances the libraries could not invest storage equipment for depositedpost

graduate thesis and dissertations CDs and therefore they were dumped in boxes with no organisation or protection. One IR had stipulated in the OA policy that 0.8 million would be set aside for among others upgrading of software as well as migration and preservation. There was however no documentation to support this.

#### **4.4.3 Technology Planning**

There were no documented plans for technology sustainability and technology monitoring as although respondents maintained that the activities were carried out. Lacking documented evidence, the evaluation could not ascertain this. There is need to have a plan outlining procedures and responsibilities for technology monitoring is not ignored until the last minute when it could be too late and users are not able to render the materials due to hardware, software format obsolescence. Frequent DSpace software updates as well as hardware maintenance were done frequently.

#### **4.4.4 Staffing and Skills Development**

Plans for skills development were contained in the overall library personnel development plans that assumed a generalized view that failed to identify skills specific to digital preservation. Repository staff attended seminars and workshops more sore tailored towards populating the repository, managing copyright, developing open access polices but so far had attended none on digital preservation. These findings reveal that although the policies of the IRs have a commitment to preserve resources under their care, digital preservation was not presently a priority as summed up by a respondent's comment below.

*“currently, we are under pressure to increase the content in the IR to improve our ranking in the next webometrics ranking.” B-DUL*

*“last time we dropped in the rankings and the VC is on one necks. For now, we are working on numbers” A-RA*

Table 17: Summary of responses on adequacy of digital preservation planning

#### 4.5 Digital Preservation Strategies

The study sought to identify the digital preservation strategies being used to protect digital resources and received responses as shown in the table below.

**Table 18: Digital Preservation Strategies Identified**

Preservation strategy	Backup	Duplication	Migration	Normalization	Refreshing	Technology monitoring	Encapsulation	Analog Copies
A	✓	✓	✓	✓	✗	✓	✓	✓
B	✓	✓	✓	✗	✗	✓	✓	✓
C	✓	✓	✓	✗	✗	✓	✓	✓

The table above shows some of the strategies that the IRs have put in place to support long-term preservation of the digital resources in their repositories. The respondents identified backup, duplication and normalization as the strategies that were actively being used.

Normalization was done through adherence to procedures during ingest, metadata creation and identification of file formats although a large 75% of these actions were not documented. For example, only one case had formalized accepted file formats via the preservation policy. The use of persistent Universal Resource Identifiers (URIs) another normalization action was common to all the IRs.

In regard to backups, the IRs backed up their data on a weekly basis. The study also established that the IRs stored as many as three copies of the same item in different locations. One IR also used a hard disk as a backup although there were no preservation actions undertaken to protect the hard disk.

Although all the universities practiced technology watch especially in relation to monitoring new updates on the DSpace software, the updates were not carried out consistently and in some cases even when scheduled they failed to take place. Monitoring of file format updates was not done.

Respondents also identified the adoption of quality control strategies at the ingest stage was a best practice that was common to all the IRs. This was collaborated by an examination of the existing policies where the following statements brought to fore this point; University A *“repository staff and quality assurance officers will review and assess all submissions before making them available,* university B *“object will be checked for valid layout and format, correct metadata and the exclusion of spam”*, university C provided a quality criterion but that was not clearly defined *“submissions not meeting the submission criteria will be returned to the depositor”*. The findings reveal that except in one instance where the criteria were clearly defined, the other IRs criteria were inadequate and were bound to create inconsistencies during implementation.

Only one IR had clearly identified migration as their preferred preservation strategy but there was no evidence that it had been applied to any the resources.

It is clear from the above findings that the digital preservation methods being utilized are not conclusive in ensuring long-term access because they only took care of risks and disasters without considering bit rot that could happen to a resource over time.

Limited knowledge on format obsolescence and its effects on digital preservation was noted from responses as indicated below:

*“so far so good, we shall deal with it when we experience it” A-ICT*  
*“I have been keeping my files for over ten years and they are still working”*  
 C-RA

Presently, the hard ware and software utilized by the three IRs were adequate but a technology plan for maintenance and replacement was lacking in all of them. Consequently, some digital preservation features provided by the DSpace software such as Checksum (an algorithm that checks for bit rot) were not utilized an issue that was linked to inadequate policies and plans that ensures that the digital resource is kept safe from all risks by assigning responsibilities as well as action plans. One IR featured dead links making some of its digital resources inaccessible because there were no strategies in place to identify these dead links for maintenance.

### **An error has occurred**

---

[Go to Repository home](#)

```
java.io.FileNotFoundException: /dspace/assetstore/97/47/64/97476411628031565248622588494217896511 (No such file or directory)
```

[Cocoon stacktrace](#) [\[show\]](#)

[Java stacktrace](#) [\[show\]](#)

[Java full stacktrace](#) [\[show\]](#)

---

The [Manakin](#) interface of the [DSpace](#) digital repository software.

**Figure 8: Sample Dead-Link Message for IR**

The use of Analog copies as a digital preservation strategy was common to all cases but was only applicable to thesis and dissertations since students by mandate were required to present both a hard copy and a soft copy before clearance for graduation. The existence of the two brought about laxity in digital preservation since the analog copy could be digitized in case of loss of the digital copy. This practice endangered born digital resources like journal articles and conference presentations that may have been received into the repository via email or when the depositor presented a copy using their own flash disk or memory card. The born digital content required more vigorous preservation measures because the IR copy may be the only one available in cases where the original owners were not in a position to provide a copy again.

The results found no evidence of documented procedures and guidelines that could be used by depositors for creation of documents, ingest (uploading) in support self-archiving although depositors were yet to start self-archiving.

#### **4.6 Challenges of Digital Preservation**

The study identified the following challenges in the preservation of digital materials: Lack of a comprehensive digital preservation policy: This was the root cause of a myriad of problems such as failure to include digital preservation in strategic plans as well as inability to perform technology planning and staff development.

Poor management of storage media. All the universities under study did not have a good storage environment for the CDs that held the original thesis and dissertations from the postgraduate directorates. When asked if this supported their commitment to long-term preservation one respondent said that “the original is always there, and we can still



digitize”. This means that the universities, did not see the need to protect the CDs since they could digitize the original thesis in case of loss of the digital copy.

Another challenge is lack of awareness on several aspects of digital preservation. These aspects include the belief that with only descriptive and rights metadata is necessary for long term access without the idea that long term access goes hand in hand with digital preservation and that technical metadata is very necessary for future generations to be able to make sense of the information by using the right software, hardware and file formats. It was also clear that some felt the issue of digital preservation was not clear to them since they thought that having a copy was enough for them.

Poor procedural accountability as evidenced by failure to document actions undertaken in the management of digital resources. Documentation indicates that the repository management is showing responsibility of any actions undertaken. This was especially felt in one of the cases where IR staff changed frequently and consistency in application of procedures was compromised.

The IR administrators in some cases are not involved in any way with the creation of the product until a copy of the CD is brought to the repository for uploading. Sometimes the CD is blank and the student has already graduated forcing them to either ignore or arrange for the digitization of the print copy.

The IRs in the selected universities are ill equipped in digital preservation in terms of skills and technology. This is because currently, there are no human resources specifically trained on digital preservation nor do they have any arrangements with other

instructions to provide them with these services. In terms of technology preparation, since there is no plan, there are no arrangements in place to replace hardware or software.

The adoption of international standards supporting digital preservation was not considered resulting in metadata schemas such as the Dublin Core metadata schema. This resulted in the creation and storage of descriptive metadata supporting access and little preservation metadata for long term preservation.

The ICT department was not conversant with digital preservation requirements as evidenced by limited procedures to check the viability of the bits as well as adherence to technology monitoring and updates.

#### **4.7 Proposed Solutions to Challenges in Digital Preservation**

The respondents gave the following solutions to the problems they encountered in regard to digital preservation:

Need for more exposure not only to issues of rights management, access, populating of the repository but also on digital preservation. In regard to poor management and storage conditions for the CDs, the respondents felt that they were not necessary and depositors needed only to provide a link to the resource unless the university was willing to invest in storage equipment for the CDs and to set money aside for media refreshing every now and then. There was also a feeling that the CDs were very important and therefore management should provide for their management and maintenance.

There is need to document policies and procedures especially in cases where there were changes in staff running the department due to transfers to other sections.

There is need to develop a comprehensive digital preservation policy to formalize the process so that a strategic plan can be made and also to ensure sustainability in terms of staff and finance for technology sustainability.

#### **4.10 Summary**

This chapter has the presented findings of the study as obtained from the interviews, observation and policy analysis regarding digital preservation in the IRs of three public universities in Kenta. The findings reveal that the IRs contain a wide range of information materials that may be used to support research by future generations. These resources are also available in a number of formats that require to be monitoring thus giving the IR managers the responsibility of ensuring the formats received in the repository are kept to a manageable number for successful preservation. They also revealed that although the IRs created metadata for the digital resources, the metadata was not enough to ensure access and use by future generations since it lacked technical metadata that informs on the resources computing environment for successful rendering of the resource.

The adoption of digital preservation planning as a best practice to successful digital preservation was minimal. The feeling that they could always go back to digitizing the hard copy also is a hindering block to fully implementation of digital preservation.

## **CHAPTER FIVE**

### **DISCUSSION OF FINDINGS**

#### **5.0 Introduction**

This chapter discusses the findings of the study whose aim was to evaluate digital preservation practices in Institutional Repositories (IRs) of selected public universities in Kenya with the aim of giving recommendations and proposing best practices for digital preservation in IRs. The discussion is guided by the following themes as identified during data analysis:

1. Characteristics of digital materials
2. Digital preservation policy
3. Digital preservation planning
4. Digital preservation strategies
5. Digital preservation challenges
6. Proposed solutions to the challenges

Data from three public universities selected purposively based on registration of IR with OpenDoar, registration of policy and mandate with ROARMAP and top three with highest content in their repositories as indicated in the repository database. Data was collected through interviews, observation and document analysis. directed content analysis was used to analyses the data and themes developed guided by ISO 14721:2012 and ISO 16363:2011.

The scope of this chapter is defined by the research questions, ISO 14721, 2012: Open Archives Information Systems framework, ISO 16363/TDR: Space data and information

transfer systems — Audit and certification of trustworthy digital repositories and the key themes and subthemes identified during the directed content analysis.

### **5.1 Characteristics of Materials Accepted in the Repository**

The research sought to find out the characteristics of resources accepted in the institutional repositories of the universities under study based on type of content, file formats, procedures for the ingest process and metadata. Content has been placed at the centre of digital preservation with many preservation decisions based on its value and level of risks its exposed to. The discussion is guided by the following themes: type of content, existence of acquisition policies that define content received, type of formats, ingest procedures as well as the metadata created to support long term access. The themes were derived from ISO 14721:2012 and ISO 16363:2011 as well as the data collection tools. The study aimed at establishing how effectively the IRs under study managed content based on these characteristics in order to support long term preservation.

#### **5.1.1 Types of Digital Materials**

There are no strict and definitive guidelines to determine the specific content types or the essential elements that an Institutional Repository (IR) should include (Roy, et al 2018). Numerous scholars have advocated for the inclusion of diverse content in Institutional Repositories (IRs) to ensure that both formal and informal modes of scholarly communication are accommodated within a unified archival platform (Robins, 2002; Genoni, et al 2004). The IRs under this study had a broad range of content that included preprints; working papers; theses and dissertations ;research and technical reports; conference proceedings and data sets among others. Primary resources like journals,

conference proceedings, thesis and dissertations formed a large part of the contents held by the IRs. of note was that some of these resources such as journal articles and conference proceedings existed in a digital format only due to the uptake of open access publishing. This is consistent with Frank (2018) who noted that repositories are entrusted with valuable digital information that necessitates the need for active digital preservation to ensure that they are available for use by future generations. However, respondents failed to make connection between type of content and long term access accepting even content not defined in their acquisition policy bringing forth the essence of Sandy and Corrado's (2017) advice to repository managers to focus on the usefulness of content within repositories in order to succeed in their digital preservation efforts. Although repositories are not excluded from collecting other types of content ISO 147211:2012 and ISO 16363: recommends that collection should be well defined and that a needs assessment be done to avoid indiscriminate ingestion of content that could lead to future resource constraints. The study also revealed a void in risk assessment for digital resources. Appraisal was also not done to identify the value of the digital resources under the custodianship of the IR. Risk assessment and appraisal are very important aspects of digital preservation because they assist the repository management to identify preservation requirements for each of the items that they have in their custody. These findings concur with Rieger, et al (2022) who decried that institutions had failed to adopt and strengthen strategies that ensured long term preservation of resources under their care.

The type of content being preserved has significant implications for long-term digital preservation. This is because different types of content require different approaches to

preservation. In addition, IRs should invest in undertake appraisal and selection activities because these actions will ensure that the resources that are being preserved are valuable, relevant, and worth the time and resources required to preserve them.

### **5.1.2 Defined boundaries for materials Accepted**

The study established that the IRs were very clear on the kind of content ingested as defined by their open access policies. However, there were instances when some of them did not adhere to the selection criteria set. More so, there was no documentation indicating how and why decisions to deviate from the set policy were made. This contradicted with ISO 14721: OAIS and ISO 16363/TRD requirement that formal selection of content deposited in the repository should be done and any action done contrary to policy documented with justification. Bearing in mind that all repositories had a commitment to preserve to perpetuity all content ingested into the IR, the need for a strong selection policy was noted. These findings collaborated with Dobratz et al (2007) who opined that the trustworthiness of a digital repository is demonstrated by its ability to operate according to its defined objectives and specifications. Identifying content boundaries through policies ensures that there are formalized decisions about what content requires short, medium or long-term preservation strategies (Anyaoku, et al, 2019). Ingestion of content indiscriminately is bound to bring challenges to the IR in future especially in relation to storage space and this is bound to test the commitment to manage content indefinitely. Weak content selection policies lead to the inclusion of low-quality or irrelevant content in a digital collection that will not only strain the resources of the IR but also be in vain when they have no value to the user.

### **5.1.3 Types of Formats Accepted in the Institutional Repositories**

The study found that to a large extent the role of file formats in long term digital preservation had been overlooked. Only one IR had made effort to minimize the number of file formats accepted in the repositories by categorically stipulating the files that will be accepted. This has been considered a good digital preservation practice that enables a repository to keep track of changes in file format and avoid obsolescence as well as reduce migration costs ((Traczyk, 2017; Margot, 2019). The use of open file formats was noted in one IR but a distribution of the actual files was lacking. The study noted a gap that would result in inconsistencies especially when there are staff changes. These findings are comparable to Francke, et al, (2017) who established that majority of the repositories under their study in Sweden had some form of guidelines for which file formats were accepted, but very few considered whether or not file formats constituted open standards.

The study identified the most common file format as the portable document format (PDF) although, MP3, Jpeg, MP4,vob and html were also used. The choice of PDF was based on its support for interoperability and therefore could support future access but studies have shown that although the standard PDF format is a good format for digital preservation, digital preservationists prefer the PDF/A as it is specifically designed for digital archiving (Corrado & Sandy, 2017). Cases of noncompliance to the recommended file formats were noted especially for video and audio formats with the repository software indicating unknown file formats. There are risks associated with ingesting unknown file formats because details about risks associated with them are unknown and could not be planned and mitigated against. The continued use of these file formats left the information resource unprotected and prone to loss in future (ZBW-Leibniz



Information Centre for Economics, n.d). failure to identify file format versions was also common in all the IRs. These findings agree with Umana (2020) who also noted that rarely did IRs identify the version of file format used. The integrity of information resources builds the trust users have on it. File formats go a long way in entrenching this trust. Understanding the format during digital preservation ensures that all features are as close to the original as possible to ensure that the user is assured that the integrity of the object is maintained. Failure to provide file version information impedes on future preservation actions leading to migration to file formats that change the outlook of the digital object causing users to question its integrity. Text-based file formats are more stable and easier to preserve than other types such as audio or video. In addition, file formats for text-based content are widely used and supported making easier to migrate to new formats.

On the other hand, audio and video file formats are more challenging to preserve, proprietary or specialized that are not widely supported. In addition, some audio and video formats require specialized hardware or software to access and play, making them more difficult to use and preserve over time. To ensure the long-term preservation of different types of file formats, the IRs need to consider the specific needs and challenges of each type of content, and develop and implement appropriate preservation strategies and policies.

Although quality control procedures described in the OA policy recommended that the IR check for the quality of format, the procedures were found to be inadequate because they lacked the criteria defining what a “quality file format” is. The findings confirm Termens,

et al (2015) observation that repository managers failed to take file formats considerations as a preservation action but this could also be attributed to limited knowledge by the respondents in regard to the relationship between file format as the study established. The unavailability of a well-defined quality control criteria when receiving files to the IR makes it difficult for staff to properly assess the quality of digital objects received to determine whether they are suitable. Quality control criteria provide clear guidelines of what is expected of the incoming materials and helps to ensure that they meet the necessary standards. This partially explains why unknown file formats are ingested into the IRs. Worse, it is bound to lead to the acceptance of materials that are damaged, incomplete, or otherwise unsuitable, which impacts on the integrity and reliability of the digital preservation efforts.

The inadequacy of the file format policy where it existed was exhibited by a lack of distinction between proprietary or nonproprietary formats. As the DPC (2015) points out, distinguishing between proprietary and nonproprietary format is an important decision that has an impact on future digital preservation activities. Proprietary formats are thought to be susceptible to upgrade and obsolescence in case the owner gets out of business. Nonproprietary offer more freedom to support digital preservation and are believed to be long-term preservation as they are believed to be stable.

Inadequate file format policies exposed the IRs to format obsolescence, lack of documentation for users and IR managers to consult in future as well as intellectual property issues during migration and emulation. It is important for the IRs to develop and implement comprehensive file format policies that take into account the long-term

preservation of digital information. This involves choosing file formats that are widely used and supported, ensuring that the necessary documentation is available, and considering the potential intellectual property issues associated with different file formats on the long term preservation of content that they manage.

Technological viability was partly achieved in the area of file format normalization by the choice of DSpace as an IR platform since it allowed the IRs to designate file formats. Respondents felt that DSpace was adequate enough to guide the IRs on types of formats because it enabled them to choose three levels of preservation format; supported, known and unsupported. Although this was considered a good practice, the functions that the platform provided were more in line with digital collection management with little much to support for long-term digital preservation. These finding mirror studied that put forward that DSpace does not have full preservation functionality, (Duranti & Shaffer (ed), 2012; Ruusalepp & Dobрева, 2012). These results point the need for the IRs invest in a platform that can support long term digital preservation because reliance on DSpace alone was a risk to the digital objects. A platform that is specifically designed for long-term digital preservation by providing tools and infrastructure that support planning, policy development and file format migration which currently DSpace cannot handle.

#### **5.1.4 Procedures for receiving (Ingesting) Materials in the IR**

Although the OA policies had a section on procedures for submission, the study established that they were inadequate to support digital preservation. They provided guidelines on who can submit and requirements for co-authors failing to provide a quality criterion on for example file format and metadata creation. In one IR, respondents reported that they were in the process of developing a guideline to support self-archiving.

This is a critical step that Baucom (2019) advises repositories to acknowledge the role of creators and in addition educate them as early as possible on technical and contextual metadata. Undefined ingest procedures posed such problems as incomplete or inaccurate metadata, data loss or corruption as well as legal and ethical risks stemming from failure to check whether items received meet the copyright threshold to the IRs.

The findings established that the IRs had no written guidelines on what constituted “quality” to guide the staff as the repositories. Instead, the IRs relied on postgraduate guidelines for thesis writing which had been developed with no consideration for preservation. ISO 16363 / TDR (2012) which proposes that trust in long term preservation should include consistence and quality management which is supported by documentation in cases where new staff maybe used to do the work. Lack of procedures impended on long term digital preservation especially in IRs that experienced a high staff turnover or in instances where casuals were employed for data entry. Pre ingest procedures such as format identification, characterization and validation, virus-checking and checksum hash generation that are critical to long term preservation were not performed due to lack of these procedures. These findings are in agreement with Langley (2019) who noted that organizations failed to understand the lifecycle requirement for digital preservation instead focusing on the object when it is received into the repository. IRs should identify content that requires long term preservation and then adopt a life cycle management approach for this content to ensure that it meets the requirements for long-term preservation. This includes following standards and best practices for file formats, metadata, and technical infrastructure.

The study established that the acquisition of both an analog and digital copy (in CDROM) of Postgraduate thesis although a major strength for the IR, poor digital media management actions exposed the CDs to environmental hazards that caused deterioration. The implications of these findings was that the IRs were receiving media formats that they could not manage. These actions also exposed the born digital resources to risk because according to the respondents, backups existed that they could turn to un case of losses. Consequently, Baucom, (2019) feel that custodians of digital resources should have policies defining how to treat digital surrogates and at the same time protect born digital resources from harm. The preservation needs of analog and digital copies may be different, and it may be necessary to develop separate strategies and processes for preserving each type of copy. The preservation needs of analog and digital copies are different and necessitates the development of separate strategies and processes for preserving each type of copy. For the IRs, this is a challenge as it involves significant costs and resources emanating from investments in technology, training, and staffing to support the preservation and access of both types of copies

### **5.1.5 Metadata**

This study sought to establish the suitability of metadata collected for long term preservation. to achieve this aim, information on metadata schema used, the types of metadata as well as the adequacy of metadata and the respondents understanding of the relationship between metadata and long-term preservation was sought.

In relation to the metadata schema, the study noted that the Dublin Core metadata scheme was common to all the IRs. All the respondents agreed that the schema was suitable for their digital preservation needs. However, these findings contrasted with Joo et al, (2019),

who termed the Dublin core metadata schema inadequate due to its limited vocabulary but was quick to add that it can be customized to accommodate comprehensive metadata. The findings of this study authenticated Joos' sentiments because gaps were identified in the types metadata collected by the IRs as discussed here.

Digital preservation metadata comprise elements such as file format file size; digital object identifiers; provenance information such as preservation actions taken on object; context information as defined by relationship of content to its environment and access right information detailing access restrictions on object. The study established that fixity information was captured in form of file size whereas representation information was minimal with file format indicated and file version missing. Reference information included a persistent URL, while provenance included the author and title. Context information was indicated as book chapter or volume number accordingly. The IRs collected rights metadata that restricted access to some digital resources to adhere to copyright laws. However, the type of license used especially for open access resources was not indicated except in once IR. Administratively there were policies guiding access especially for copyrighted material as well as policies defining different stakeholders and what each of them can perform on a digital object. These policies ensured that any changes performed on a digital object was captured in metadata in order to protect the integrity of the digital object. OAIS specifically requires metadata in the form of Preservation Description Information (PDI), which should include provenance, reference, fixity, contextual, and access rights information, all which contributes to maintaining a digital object's authenticity, (ISO 14721, 2012). Although there is a dearth of research, usage, quality and practices of preservation metadata in the IR context (Lee & Stvilia,

2017; Termenes, 2022), these findings correlate with Magama (2017) sentiments that custodians of digital records did not collect enough preservation metadata to support the long-term preservation in Masvingo province, Zimbabwe.

Gaps in the collection of preservation metadata was attributed to limited knowledge on the role of metadata in long-term preservation. These findings mirror Magama (2017) who noted a general ignorance on the role of preservation metadata in the management of digital records recommending training to raise awareness about the same. Preservation metadata supports emulation and migration to avoid technology and format obsolescence. Although descriptive metadata is important for providing context and information about digital objects, such as documents, images, and audio files relying solely on it can have some catastrophic effects on long-term digital preservation. This is because descriptive metadata does not provide information about the technical characteristics of the object, such as the file format, size, and resolution. It also does not provide information about the structural relationships between different objects, or about the administrative processes involved in managing the object over time. To ensure the long-term preservation of digital objects, the IRs should also include technical, structural, and administrative metadata. Technical metadata helps with understanding and preserving the technical characteristics of the object, while structural metadata helps with understanding the relationships between different objects. Administrative metadata helps with managing and preserving the object over time, including information about rights management, access control, and preservation actions.

To address gaps in metadata, it is important for the IRs to invest in robust metadata creation and management processes, and to prioritize the ongoing maintenance and update of metadata over time. This may involve training staff and other stakeholders on metadata standards and best practices, and establishing processes for regularly reviewing and updating metadata to ensure its accuracy and completeness.

## **5.2 Preservation Policy**

Ismail and Affandy (2018) described policy development as one of the most vital digital preservation strategies as it stipulated what needed to be preserved. The findings of this study established that although the IRs had a section in their OA policy, it was not adequate enough since it lacked crucial sections such well-defined roles, a description of the digital preservation context, a description of the characteristics of the collection as well as identification of digital preservation strategies. Barrueco and Termens (2022) came to the same conclusion that although IRs in Africa had policy statements, they could not be equated to a proper policy. Da Silva and Borges (2017) in a study within Brazilian IRs noted similar cases where repositories committed to long term preservation but failed to have programmes to guide this commitment. Policy statements set clear priorities, and ensure stakeholder collaborations by making it the responsibility of the organization to safe guard its digital resources for the future generations. Policy also ensures that the organization defines organizational roles and funding responsibilities. The consequences of missing or inadequate policies exposed the IRs to poor strategic planning and poor funding and the adoption of passive digital preservation where digital objectives requiring long term preservation are not identified and taken care of even before creation.



Since digital preservation was incorporated as a section of the OA policy, the review of the OA policy was considered to establish whether a review of the preservation section was carried out. One IR did not indicate a review date while the other two capped the review cycle at 5 years. Never the less, none of the IRs had carried out any review. Respondents were quick to point out that the IRs were still in infancy and that fear of failure was the main reason that they could not commit to review the policy. The findings correspond to Dressler (2017) findings that digital preservation policies of ARL member libraries lacked a section on review and instances where this had been given three years had passed since the review had been done. This she termed as inaction and the worst form of response by caretakers of digital objects. Failure to review policy implied that the IRs were not keen to respond to changes in technology by capturing these changes in their policies. This exposed the digital resources under their care to many risks among them technology obsolescence. Failure to frequently review the policy meant that changes in user needs could also not be implemented and resources that met these new needs were exposed to risks.

There were no clearly defined roles and responsibilities for digital preservation stemming from a limited digital preservation policy. Defining roles and responsibilities helps to ensure that there is clarity about who is responsible for different aspects of the digital preservation process, which can promote efficiency and effectiveness.

The policies were very clear on copyright aspects since depositors were required to sign deposit agreements granting repository staff permission to store, copy and format/manipulate the materials in order to ensure that it can be preserved and made

available in the future. ISO 14721:2012, OAIS Reference model requires that the digital repositories negotiate for their content from the creators. Hoeren et al (2013) argued that the exclusive rights given by copyright such as right to copy and right to alter/modify were very crucial to digital preservation necessitating the need for IRs to negotiate for digital content from the creators in order to be able to undertake preservation actions without seeking permissions from them. Negotiating for copyright was the right step in digital preservation because it enabled the IRs to undertake preservation actions without consulting the content owners. Seeking permissions is sometimes time consuming and may cause loss of a digital object before permission is granted.

### **5.3 Digital Preservation Plan**

A digital preservation plan defines and documents the vision and strategy of long-term digital preservation as well as all the important definitions that will make the implementation of the digital preservation policy accurate and complete (Bountouri, 2018).

This study established that all the IRs under study did not have a documented preservation plan. The lack of a plan stemmed from a lack of a comprehensive digital preservation policy. The findings also reveal a general lack of awareness of preservation planning among the respondents these findings are consistent with IRMT (2016) who noted a general lack of awareness across stakeholders in regard to digital preservation planning. These findings have a number of implications on the IRs. Without a plan in place, it is difficult or impossible for the IR to identify and prioritize its preservation needs, allocate the necessary resources and expertise as well as implement appropriate

preservation strategies. Lack of planning also leads to inefficiencies and waste especially of storage space. The IRs commit to preserve all content ingested to perpetuity a requirement that may find them preserving a lot of content with little value. In addition, The IRs are likely to incur financial losses to correct errors that would not have occurred with planning.

Due to decreasing budgets, none of the IRs had set any funds aside for preservation. According to the respondents, the funds the IRs received were directed towards increasing content and providing access. These findings agree with other studies that have found that IRs seem to focus on acquisition of content at the expense of preservation due to dwindling budgets (Adjei, et al, 2019; Robertson & Borchert, 2014; Wittenberg, 2018). ARMA (2015) recognizes the importance of financial sustainability by arguing that long term digital preservation is expensive not only due to the infrastructure required but also funds for staffing and technology watch and therefore an institution committed to digital preservation must be able to demonstrate long term financial viability. Financial sustainability impact on digital preservation efforts in a number of ways such as the costs of acquiring and maintaining the necessary technology, the costs of training and staffing, and the costs of ongoing preservation activities such as migration and emulation. The lack of funding noticed in the IRs meant that the IRs cannot undertake any of these actions.

The study found the IRs had currently invested in hard ware and software with all three cases utilizing the DSpace platform. Being an open source software, DSpace allowed for customization, scalability and enabled the IRs to manage versions as well as check bit rot

through its checksum function. It also enabled the IRS to manage a number of file formats as well as diverse content. However, the software lacked functionalities to support long term digital preservation actions such as file forward migrations. These findings agree with Dowding, (2016) who noted that DSpace as a long-term digital preservation platform lacked some long term digital preservation functionalities and in addition failed implement metadata standards like PREMIS that could enable it to collect enough metadata to support digital preservation. PREMIS is designed to be flexible and extensible, allowing digital repositories to tailor it to their specific needs and requirements. It covers a wide range of metadata elements, including technical, administrative, and intellectual property information, as well as information about the relationships and dependencies between different resources all critical to successful digital preservation.

Although the IRs had good technology infrastructure they however fell short on technology planning as they all lacked a policy and plan that stipulated future directions that will be taken in regard to maintenance and replacement of the technology. McGovern and McKay (2008) acknowledged that technology planning was very critical in digital preservation as it assisted organizations to anticipate needs as well plan for infrastructure with the full support of top level management. Poor technology leads to the use of inappropriate technology. For example, currently the IRs are using DSpace that lacks some functionalities to support long term preservation. Additionally, since no funds have been set aside for upgrade or maintenance the IRs risk finding themselves with obsolete technology. This study concluded that poor technology planning exposed the IRs serious consequences that were bound to hamper their digital preservation efforts.

There were no strategies identified to develop skills critical specific to digital preservation. Rather, staff development plans were provided for the whole library. Repository staff attended seminars and workshops mostly on open access and licensing but none related to digital preservation. This agrees with Umana, (2017) who noted that little was being done to equip staff with digital preservation skills although it was highly technical and complex. There have been recommendations for libraries to invest in skills development and up skill the IR staff through workshops and seminars (Masenya & Ngulube, 2019) as well through information exchange (Alison, et al 2019) in order for IRs to meet this need. Without adequate planning and development of skills, it was difficulty to ensure that there were enough trained and knowledgeable staff to properly manage and preserve digital resources. This often leads to a lack of efficiency and effectiveness in the preservation process, as well as an increased risk of errors or omissions that could compromise the integrity of the preserved materials. Poor skills planning and development also hinders the ability to adapt to new technologies or methods of preservation, leading to a reliance on outdated techniques.

#### **5.4 Digital Preservation Strategies**

Digital preservation strategies encompass various approaches that ensure the long-term safety, accessibility, and usability of digital objects. These strategies play a crucial role in safeguarding digital assets for as long as they are needed. The study identified several key strategies were namely backup (or bit stream copying), replication, and the adoption of procedures and standards.

Backup, commonly employed by IRs, involves creating copies of digital objects to mitigate the risk of loss. It is widely recognized as a short-term remedy rather than a long-term preservation solution ((Moseti, 2016; Adjei, et al 2019; Srirahayu, et al, 2020). Scholars and experts have cautioned against relying solely on backup strategies, emphasizing their limitations in ensuring sustainable digital preservation. UNESCO and other studies have stressed the need for additional measures, as relying solely on backups may leave repositories without a viable copy ((UNESCO, 2003; Srirahayu, et al, 2020). Regular checks for bit rot (data degradation) through checksum runs and broken link detection are recommended but were found to be lacking in the surveyed IRs.

The study established that replication was achieved by having many copies distributed in many sites as well as use of external media such as a hard disk. Replication is regarded as a critical component of digital preservation, significantly reducing the risk of loss (Gallinger et al, 2017). However, it introduces new challenges as each replicated copy must undergo its own preservation processes to maintain long-term usability (Magama, 2017).

The adoption of procedures for file formats, ingest (normalization) and the use of Dublin Core metadata schema (use of standards) were the major long-term digital preservation strategies. File format procedures prove useful in limiting the number of formats ingested into a repository, facilitating change monitoring, and reducing costs associated with format migration (Traczyk, 2017, Margot, 2019). By adopting standardized file formats, IRs can ensure compatibility, interoperability, and long-term accessibility of digital

objects. However, it is important for repositories to consider the evolving nature of file formats and stay updated on emerging standards.

In addition to file format procedures, ingest (normalization) procedures and the use of metadata schemas, such as the Dublin Core, are critical long-term preservation strategies. Ingest procedures involve the transformation of digital objects into a standardized format for better preservation and future access. The adoption of the Dublin Core metadata schema, while common among IRs, has limitations due to its limited vocabulary. As digital preservation requires a wide range of metadata, repositories need to consider more robust and comprehensive metadata standards that accommodate diverse preservation requirements.

The study concluded that the digital preservation strategies used by the IRs were short term that were not adequate to protect born digital resources that if lost could not be replace. These are in tandem with Umana (2019) who noted that IRs tended to rely on short term digital preservation strategies endangering the resources under their care. No single preservation strategy can work alone and digital repositories have been advised to adopt both short- and long-term preservation strategies in order to ensure digital resources in their possession survived technological changes, (Vrana, 2011).

Short-term strategies cannot adequately address the challenges and risks associated with preserving digital content in the long term. Short-term strategies fail to consider file format, hardware and software obsolescence that if they left unconsidered exposes the content in the IRs to inaccessibility in future.

## **5.5 Challenges of Digital Preservation**

The study identified the following challenges in the preservation of digital materials: Poor management of storage media. Storage environment for the CD ROMs that held the original thesis and dissertations from the postgraduate directorates was inadequate as they were placed in boxes and stored in dusty rooms. This contravenes Bacciu et al, (2019) notion that digital repositories have a responsibility to protect any digital media in their custody from deterioration and obsolescence.

Another challenge was the lack of awareness on several aspects of digital preservation. These aspects include the belief that only descriptive and rights metadata were necessary for long term access. There was limited knowledge on the relationship between long term access and digital preservation. This resulted in little attention to the creation, capture and maintenance of technical metadata and more reliance on descriptive metadata. The results deviated from Masenya and Ngulube (2020) who reported an increase in awareness on the consequences of poor preservation actions in South Africa.

Another challenge was failure to document actions undertaken in the management of digital resources. Documentation indicates that the repository management is showing responsibility of any actions undertaken. This was especially felt in one of the cases where IR staff changed frequently and consistency in application of procedures was compromised. ISO 16363/TDR requires that organizations document all repository processes, decisions and goal setting so that these activities may be understood by stakeholders. The IRs should demonstrate sustainability by adequate procedural accountability since these procedures can be consulted in the future when one wants to



know why, what, or how decisions were made and what actions were taken (Lampert & Vaughan 2018).

The IR administrators were not involved in any way with the creation of the product until a copy of the CD ROM was brought to the repository for uploading. Sometimes the CD ROM was blank and the student had already graduated forcing the IR staff to either ignore or arrange for the digitization of the analog copy. IRs should embrace the life cycle management of digital content and work with the creators of the content and avoid waiting until the files are ingested into the repository. This is a problem that can be solved through the life cycle approach as discussed in an earlier paragraph.

The IR staff were also ill equipped for digital preservation in terms of skills and technology. There was nobody specifically trained on digital preservation nor were there arrangements with institutions with the skills to provide them. Cunningham, (2018) proposes collaborations to help repositories cope with the challenges of staff development. By partnering with other institutions, joining professional communities, working with experts, and participating in training programs, the IR staff will gain the knowledge and skills needed in digital preservation

### **5.6 Proposed Solutions to Challenges in Digital Preservation**

The respondents gave the following solutions to the problems they encountered in regard to digital preservation:

There was need for more exposure not only on issues of rights management, access, populating of the repository but also on digital preservation especially through staff

training. Limited digital preservation knowledge and skills among staff have been identified as the greatest threat to digital preservation (Masenya & Ngulube, 2019). Kavishe (2016) recommends that institutions should invest in the re-skilling of the staff involved with digital repositories to secure institutional memories. The respondents also proposed the documentation of policies and procedures especially in cases where staff changes were common due to transfers to other sections.

To eliminate CDROMS the IR staff recommended that creators submit content by creating and providing the link to the IR. One IR had already started receiving the content through email. These findings are in agreement with Nieh et al (2018) who concurred that libraries were concerned on the viability of the CD-ROM and advised them to alternatives to the storage media to safeguard access to the resources they hold from an uncertain future. Although these were good solutions to manage storage media, they introduced other challenges for the IR to maintain the link and email for use as a backup in case the copy in the repository is lost or spoiled.

Development of a comprehensive digital preservation policy to formalize the process to facilitate the inclusion of digital preservation in the strategic plan was also proposed. This was viewed as a way to ensure sustainability in terms of staff and finance for technology planning. Policies compliance by various stakeholders as well as provide cohesion and adoption of best practices in the organization, (British Library, 2013; Dressler, 2017).

## **5.7 Summary**

This chapter summarizes the findings of this study. The findings were discussed relative to other researcher's findings and summarized as follows. The study identifies the

absence of clear guidelines for content types lack of a defined link between content type and long-term access inadequate collection policies and insufficient risk assessment and appraisal practices as well as a significant oversight regarding the importance of file formats in long-term digital preservation within Institutional Repositories (IRs). In addition, Portable Document Format (PDF) emerged as the most prevalent format and noncompliance with recommended file formats for video and audio formats. The lack of a documented ingest and metadata policies resulted in incomplete and inaccurate metadata that was inadequate for long term preservation. Policies and plans were inadequate with short term digital preservation practices being adopted.

## CHAPTER SIX

### SUMMARY, CONCLUSION AND RECOMMENDATIONS

#### 6.0 Introduction

This chapter looks presents a summary of the findings, discusses the findings, draws conclusions from the findings and gives recommendations based on the conclusions drawn. The findings are guided by the research questions.

#### 6.1 Summary of Findings

##### *Characteristics of materials in the IRs*

The study sought to determine the types of materials accepted in the IRs. The research found that the contents of the institutional repositories being studied were diverse, with 78% being primary sources. Theses and dissertations were available in both analog and digital copies, while open access research articles, journal articles, and conference proceedings were only available in digital form.

Having a hybrid collection, which includes both digital and analog copies of materials, reduced the need to invest in digital preservation efforts because the risk of losing access to the information is mitigated by the existence of the analog copy, which can be digitized if necessary. At the same time the existence of the two copies increase the resource requirements for the IR leading to the negligence of digital preservation media.

The study established that the digital materials being added to the institutional repositories were not always in line with the established acquisition criteria. The lack of a comprehensive acquisition policy made it difficult for the repositories to develop

effective appraisal strategies to identify materials for long-term preservation. These findings point out that although the institutional repositories had declared their intention to preserve the resources in their care indefinitely, they did not have concrete plans in place to ensure this was achieved.

The research found that the file formats accepted in the institutional repositories were clearly defined, with the portable document format (PDF) being commonly used for text documents, JPEG for images, MP3 for audio, and MP4 and VOB for videos. However, there were instances where the repositories received materials in formats that were not defined by their policies.

The results showed that long-term preservation was not the main consideration in the choice of file formats, and the repositories did not consistently adopt preservation-friendly, non-proprietary formats. The use of DSpace as a repository software allowed the repositories to categorize file formats as known, supported, and unknown, and provided other digital preservation features such as checksum for bit rot, but neither the repository staff nor the IT personnel were aware of how to use these features.

There was a general lack of awareness among the respondents about the importance of format management for long-term digital preservation, as evidenced by instances where photographs were converted to the PDF format. The study also highlighted the need for a comprehensive digital preservation policy outlining the file formats that would help the repositories achieve their goal of long-term digital preservation as outlined in their open access policies.

Guidelines on how resources will be received in the repository were verbal and passed on from one staff to the other. There were no written guidelines from the IRs on what the requirements for ingest and metadata collection were. For ingest, the IRs relied on guideless from the graduate school for thesis and for those that have been published, the publisher guidelines.

In relation to metadata, the Dublin Core metadata schema which is a general metadata standard was used. Limited knowledge on digital preservation specific metadata standards such as Preservation Metadata: Implementation Strategies (PREMIS) was also evidenced.

Metadata collected was inadequate to support long term digital preservation. The IRs focused more on collecting descriptive metadata to facilitate access. Technical, structural and administrative metadata were also very important for long-term digital preservation because in addition to describing the computing environment, they showed relationships between objects as well as recorded the management actions required to maintain the digital object

The study established that respondents failed to understand the need for extra metadata for digital objects. This misunderstanding stemmed from the librarian's roles in cataloguing necessitating the need to raise awareness among libraries on the need for extra metadata to support digital preservation.

### ***Digital preservation policy***

The research sought to find out the suitability of the digital preservation policy to support long term preservation.

The study found that the IRs had a digital preservation policy section within their OA that retaliated their commitment to preserve but lacked the comprehensiveness to support long term preservation. A digital preservation commitment requires an articulation of actions that the IR will undertake to actualize and sustain digital preservation.

There were two IR policies that were reviewed, but neither had a specific time frame for review. This was a mistake, as changes in user needs and technology are important factors that should be considered when reviewing policies. Efforts were made to obtain resources for the IRs through mandates outlined in the open access policies, which required depositors to grant permissions to IR administrators for preservation actions such as backup, replication, emulation, and migration. These actions can involve the infringement of creators' rights, so their cooperation is necessary for long-term digital preservation.

### ***Digital preservation planning***

The research sought to find out the adequacy of digital preservation planning to support long term digital preservation.

The study found that the IRs did not have a documented plan for preservation, but they were willing to develop one as a guide for future preservation efforts. The lack of a comprehensive preservation policy was identified as a contributing factor to the lack of

planning, as such policies provide guidance and frameworks for planning. The study also revealed that the IRs' digital preservation program was reactive rather than proactive. This lack of planning meant that the IRs were unable to identify their preservation needs and seek external help when those needs could not be met in-house. Poor planning also resulted in a lack of identification of actions for long-term digital preservation explaining why the IRs relied on short term actions such as back up and replication. Despite being in operation for eight years, and using short term digital preservation actions, the respondents felt that these measures were sufficient enough to long-term preservation. This is a significant issue, as the IRs manage digital resources that are dependent on technology, which is rapidly changing and can become outdated in a short period of time.

The research showed that while the IRs claimed to prioritize preservation, they had not yet treated digital preservation as a strategic priority. This was reflected in their budgeting, which focused on digitization and marketing rather than preserving digital media. Additionally, the CD ROMs used for preservation were being stored in poor conditions, increasing the risk of deterioration.

The study found that the IRs were satisfied with DSpace as a preservation platform. This was attributed to the fact that DSpace was an open source software that could be customized. In addition, it was scalable and provided support for a wide range of file formats, including text, audio, video, and images. Other features that endeared DSpace to the IRs are its support for versioning, which allows users to track changes to digital objects over time, checksum validation, which helps to ensure the integrity of digital objects by calculating a unique code for each object and comparing it to a stored value.



However, its major weakness was its failure to support sufficient metadata for long term preservation

The research established that there were no plans in place for the development of digital preservation skills. Although IR staff attended seminars and workshops, they focused open access, populating and marketing the IRs and not on digital preservation.

### *Digital preservation strategies*

The study found several digital preservation strategies being utilized. These included: backup, duplication, migration, normalization, refreshing, technology monitoring, encapsulation and keeping of analog Copies. Backup was done at least on a weekly basis. Duplication involved use of more than one server, offsite storage and storage media. Normalization was used to a limited extent through the use of standards such as the Dublin core metadata schema. Technology watch was also being utilized to keep track of DSpace software updates but the same was not extended to file formats. No documentation was available in form of a technology watch report to support this.

The above findings point to inadequate digital preservation actions that may not conclusively protect digital resources for long-term access because they are short term measures that only took care of risks and disasters. In essence the IRs relied on backups and multiple site storage without performing checksums to check for bit rot. Link rot check was also not done. Digital preservation being a series of actions requires more than just backup even in cases of physical media management.

### *Challenges to digital preservation*

The study uncovered several challenges associated with the preservation of digital materials. These challenges include inadequate management and storage practices for storage media, a lack of awareness regarding various aspects of digital preservation, a failure to document actions taken in the management of digital resources, and limited quality control measures. The evidence suggests that administrators of Institutional Repositories (IRs) did not provide guidelines on content creation until the final product was ready for ingestion. Furthermore, the selected universities' IRs were found to be inadequately equipped in terms of both skills and technology for digital preservation.

Regarding technology preparedness, the study revealed that there was no comprehensive plan in place for the replacement of hardware or software. This lack of planning poses a significant risk to the long-term preservation of digital materials. Without appropriate arrangements and provisions for technology upgrades or replacements, the IRs may encounter difficulties in maintaining the accessibility and usability of their digital resources.

The identified challenges in digital preservation highlight critical areas where improvements are needed. It is imperative for institutions to address the poor management and storage practices of storage media, increase awareness among stakeholders regarding digital preservation, document actions taken in the management of digital resources, and implement robust quality control measures. Additionally, investments in acquiring the necessary skills and technologies for digital preservation are crucial to ensure the long-term viability and accessibility of digital materials. Establishing a comprehensive plan for technology replacement or upgrades is also essential to mitigate

risks associated with outdated hardware and software. By addressing these challenges, institutions can enhance their digital preservation efforts and safeguard their valuable digital assets for future generations

***Proposed solutions to digital preservation challenges***

According to the respondents, there is a pressing need for increased exposure and knowledge not only regarding rights management, access, and populating of the repository but also regarding digital preservation. They expressed the importance of implementing policies that would allow students to provide links to their theses and dissertations, thus eliminating the challenges associated with managing CD-ROMs. Furthermore, the respondents emphasized the significance of documenting policies and procedures, particularly in cases where staff members responsible for managing Institutional Repositories (IRs) may transfer to other departments.

Another crucial concern raised by the respondents was the necessity of developing a comprehensive digital preservation policy. They emphasized the need to formalize the preservation process, enabling the creation of a strategic plan. Such a policy would also address concerns about the long-term sustainability of the IRs, including aspects such as staffing and financial support for technology maintenance.

By acknowledging the need for additional exposure and understanding, institutions can equip their staff and stakeholders with the knowledge necessary for effective management of digital repositories. Implementing policies that encourage students to provide online links to their work can alleviate challenges related to physical storage media. Furthermore, documenting policies and procedures ensures continuity and ease of

management, even in the event of staff transitions. Lastly, the development of a comprehensive digital preservation policy provides a framework for strategic planning and sustainability, ensuring that the IRs can continue to fulfill their mission in the long run.

Addressing these concerns and implementing the suggested measures will contribute to the improvement of digital preservation practices, fostering a more efficient and sustainable approach to managing and preserving digital materials within the IRs.

## **6.2 Conclusion**

Successful digital preservation programs are a crucial requirement in IRs to support long term accessibility of the digital resources in their custody. Institutional repositories in Kenya have made a commitment to collect, store, preserve and disseminate the resources in their repositories but little is being done to realize the same. Best practices have not incorporated in digital preservation as evidences by: lack of a comprehensive digital preservation policies and plans; content selection policies, procedures for content creators, technology plans as well as staff development strategies to equip staff with digital preservation skills necessary to protect digital resources. It was also evident that the staff in the IRs had limited knowledge on what digital preservation encompasses. Based on this then, the study drew the conclusion that the digital preservation practices in place within the three IRs could not guarantee long term preservation of the resources in the IRs due to limited adoption of best practices as well as reliance on short term digital preservation strategies.

### 6.3 Recommendations

The study has several implications to the university libraries that have established and are establishing institutional repositories and the industry as a whole.

1. The repository management should undertake a comprehensive evaluation of the institutional repository's objectives, incorporating a thorough needs analysis and a comprehensive inventory of the digital collection. Based on the findings from these assessments, the management should formulate specific and measurable goals that align with the repository's purpose and mission.
2. The policy framework should be meticulously developed to align with the identified goals. The policy must explicitly define the scope of digital preservation, whether it's short-term, medium-term, or long-term. To ensure a comprehensive and effective policy, it should be a collaborative effort involving archivists, information technologists, librarians, and university management. Additionally, the acquisition and selection policies should be robust, connecting them to the universities' overarching mission and vision. This integration will facilitate formalization and inclusion in strategic plans, considering the sustained financial commitment required for digital preservation, encompassing technology planning and expertise development
3. It is imperative to integrate digital preservation into the university's comprehensive strategic plan, guaranteeing allocation of necessary budgets. This inclusion will address issues like the absence of skill development opportunities in digital preservation that result from inadequate planning. Furthermore, it will encompass technology monitoring and planning. The university library should

assume an advocacy role to raise awareness among university management and construct a compelling business case for digital preservation, demonstrating its value and significance.

4. The IRs need to invest in research to establish the best way to go about digital preservation. Because of the high rate of technology obsolescence, it may become necessary to form collaborations to share resources not only of infrastructure but expertise to support digital preservation programs within the country.
5. Digital resources require a life cycle approach to preservation and therefore the library management should put structures in place to achieve this by for example developing procedures and templates to guide different stakeholders in areas such as creation, metadata creation and storage.
6. Just as KLISC is taking initiative to promote the open access publishing model for scholarly resources, it should take initiative to create awareness on the importance of digital preservation. This can be achieved by benchmarking with digital preservation coalitions in other parts of the world in order to gain and borrow expertise for implementing strategies identified for the Kenyan IRs.
7. For university management, it is important to appreciate that though the IRs are a means to improve the visibility of the university and at the same time make the research available to their researchers, it comes with a responsibility ensuring that access and use can be sustained in the long-term. This requires an understanding of the unique characteristics of digital resources as compared to the physical resources such as their machine dependency, format dependency, media fragility, and technology dependency. Understanding these unique attributes facilitates the

adoption of appropriate actions to preservation them. Repository administrators should up their advocacy skills through workshops and case studies focusing on university senior management in order to help them understand more.

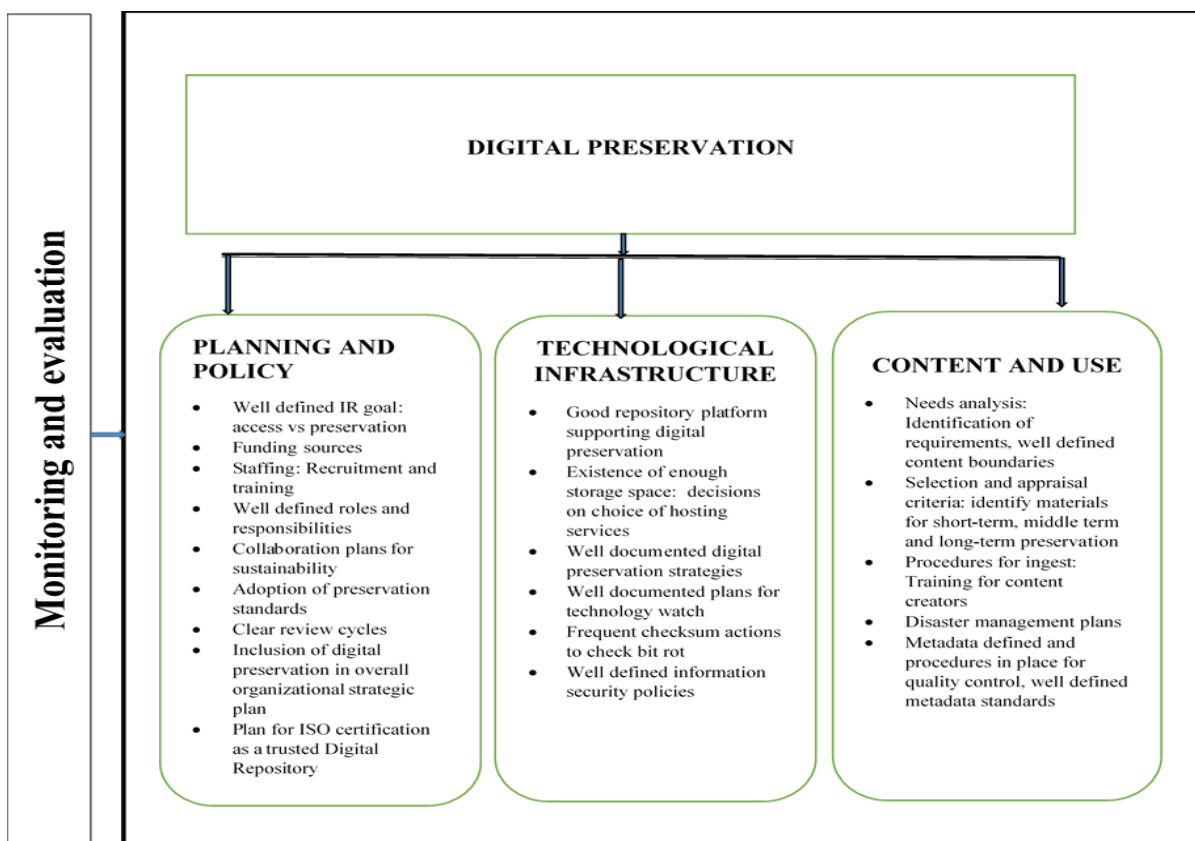
8. Just as the universities pursue certification in other areas of quality through ISO certification, they should also pursue certification for the repositories to ensure they are at par with other research institutions in the world. This can only be achieved by putting measures in place to ensure that they qualify to be trusted repositories by providing policies and plans on digital preservation, providing technological infrastructure for digital preservation, providing policies for content and access. There is need to adopt international digital preservation standards like the OAIS model and ISO 16363 Trusted Digital Standard in order to be at par with the rest of the world and to be able to fulfil one of their objectives supporting interoperability. Building a business case would be the best action to assist the repository administrators to achieve this by linking trust by researchers to their depositing research outputs to attraction of research grants.
9. Formation of digital preservation collaborations and cooperatives: this is because digital preservation is complex and at times could be beyond one institutions capability. Collaboration in either skills development, policy development, planning, standards development or technological infrastructure would go a long way into improving the current state of affairs. Cooperation enhances the productive capacity of a limited supply of digital preservation funds, by building shared resources, eliminating redundancies, and exploiting economies of scale.

10. **Reliable funding:** Ensuring reliable funding is crucial for the long-term sustainability of Institutional Repositories (IRs). To achieve financial soundness, IRs should develop comprehensive business plans and adhere to transparent business practices. The business plan should outline strategies for generating income to cover operational costs and address contingencies effectively. The business plan serves as a valuable tool for seeking funds and grants dedicated to supporting digital preservation efforts. Additionally, the business plan can facilitate the development of an economic model that ensures the IRs have continuous access to sufficient resources to meet their digital preservation objectives. By prioritizing financial stability and strategic planning, IRs can enhance their capacity to sustainably preserve and manage digital content.
11. To enhance digital preservation within Institutional Repositories (IRs), it's crucial for the IT department to gain a deep understanding of the specific preservation requirements for digital resources hosted in the IRs. This understanding should encompass the necessity for routine checks on digital items and a willingness to be flexible when certain preservation tools deviate from the organization's established policies. To implement this recommendation effectively, it's advisable to foster a robust and collaborative relationship between the digital preservation team and the IT department. Designating a dedicated member within the IT team who specializes in preservation can significantly contribute to aligning IT practices with preservation goals.



## 6.4 Proposed Digital Preservation Best Practices Framework for Institutional Repositories

The study identifies four components of an effective digital preservation program in institutional repositories. These are planning and policy; technological infrastructure, content and use and monitoring and evaluation. Figure 11 summarizes these components and is followed by their description.



*Figure 9: Proposed Framework for Best Practices for Digital Preservation*

### Planning and Policy

In order to establish a robust and effective preservation component within an Institutional Repository (IR), several key factors should be addressed.

Of critical importance is the need for the IR's mission and mandate to explicitly specify a specific goal for the repository, particularly emphasizing its commitment to long-term preservation. By clearly defining this objective, the repository can align its efforts and resources towards achieving sustained preservation of digital content.

A crucial element of the preservation component is the development of a comprehensive digital preservation policy. This policy should encompass various aspects, including a justification for preservation, its relationship with other policies within the institution, organizational and financial commitments to preservation efforts, guidelines for metadata creation, delineation of roles and responsibilities, and provisions for staff training and education. It is imperative for the policy to have a clear review cycle that accounts for changes in technology and evolving user needs. This review cycle should be flexible enough to incorporate updates or adaptations based on reports from technology watch initiatives.

Furthermore, it is essential for the IR to establish a comprehensive plan for the preservation of digital content. This plan should address key elements such as identifying the relevant organization or repository, specifying the current status of the preservation plan (whether approved or not), providing a detailed description of the institutional context, outlining the nature of the collection (including its digital objects), clearly defining the purpose and requirements of the preservation effort, presenting evidence of the decision-making process for specific preservation actions, considering cost considerations, establishing triggers for periodic re-evaluation of the plan, defining the roles and responsibilities of involved individuals, and ensuring the availability of well-

documented preservation workflows. This holistic approach ensures that all necessary aspects of preservation are considered and properly documented.

The preservation component should also prioritize training and development initiatives to sensitize all stakeholders about their roles and responsibilities in digital preservation. This includes identifying the skills required for effective preservation management and devising a staff recruitment and development plan that addresses these needs.

Fostering cooperation and collaboration with like-minded institutions is vital for long-term digital preservation. Recognizing that no single organization can solely accomplish the task, the IR should actively engage in partnerships and alliances with other institutions that share the same preservation goals. By joining forces, sharing knowledge, and pooling resources, these collaborations enhance the collective efforts towards sustainable preservation practices.

By addressing these factors and implementing the necessary measures, an IR can establish a comprehensive preservation component that ensures the long-term viability and accessibility of its digital collections.

### **Content and Use**

In order to effectively address this component, Institutional Repositories (IRs) should undertake the following actions:

Conduct a user needs assessment to identify the specific requirements for content types, software and hardware, as well as skills needed. This assessment plays a crucial role in shaping digital preservation decisions. Moreover, conducting a needs assessment brings together various stakeholders and fosters a sense of unity in achieving preservation goals.

The outcome of the needs assessment should result in a classification system for content, categorizing resources into short-term, mid-term, and long-term preservation requirements.

Collaborate with experts, including archivists, to develop a selection and appraisal policy. This policy helps identify materials suitable for short-term, mid-term, and long-term preservation. Collaborating with experts ensures that the policy aligns with industry best practices and standards.

Develop guidelines for depositors based on the results of the appraisal process. These guidelines should assist content creators in familiarizing themselves with the repository's policy, understanding the typology of digital content, selecting appropriate file formats, and fulfilling metadata requirements prior to submission. For content designated as long-term, archival file formats should be recommended to ensure their preservation over time.

Establish procedures to verify the quality of items ingested into the repository. This includes activities such as fixity checks, virus scans, normalization of files, and content packaging. Clearly outlining these procedures helps ensure effective quality management throughout the preservation process.

Develop a metadata requirements guideline that outlines the purpose, types, and elements of metadata to be collected. It is crucial to select appropriate metadata schemas that support long-term digital preservation, especially for format-specific standards. Adequate metadata collection is essential to support the ongoing preservation and accessibility of digital content within the IR.

Implement disaster management planning and ensure that staff members are appropriately trained. Having a comprehensive disaster management plan in place helps mitigate potential risks and ensures the repository is prepared to handle unexpected events. Staff training is crucial to effectively respond to disasters and minimize potential damage to the digital collections.

By implementing these measures, IRs can establish robust preservation practices, aligning their operations with industry standards and best practices. This helps to ensure the long-term preservation and accessibility of digital content within the repository.

### **Technological Infrastructure**

Technological infrastructure is very critical in the establishment of effective digital preservation practices within the (IRs). In order to achieve this, several important considerations and actions should be undertaken.

First, it is crucial to select an IR platform that not only supports digital preservation but also allows for customization to adhere to standards in this field. As most IRs currently utilize DSpace, customization should be prioritized to enable the capture of preservation metadata. This customization ensures that the necessary information is recorded and preserved alongside the digital objects.

Ensuring sufficient storage capacity is essential, and plans should be developed to meet future storage needs. This requires conducting a needs analysis and establishing selection and appraisal policies. The chosen systems should have reliable support, whether through documentation, community support, or vendor support where applicable. Scalability should also be considered to accommodate the growing volume of digital content.

The choice of hosting services for the IR's server is another critical aspect that should be backed by evidence. This decision may involve options such as hosting solely by the library, sharing a server with the institution, or utilizing a commercial host. The factors of scalability and reliability should guide this choice, and if a third-party host is selected, the contractual agreement should clearly outline the digital preservation actions to be taken.

Collaboration between the IR and the ICT department is essential to develop comprehensive policies for information security. These policies ensure the integrity and availability of the resources. Currently, policy development is often left solely to the ICT department, without input from repository administrators or systems librarians. Involving all relevant stakeholders in policy development strengthens the overall security and preservation efforts of the IR.

Active technology monitoring is necessary to stay informed about changes in technology and file sizes that may trigger preservation actions. This includes being aware of developments that may necessitate format migration, platform migration, or changes in hardware or software. By keeping a close watch on technological advancements, the IR can proactively adapt its preservation strategies to ensure ongoing accessibility and usability of digital content.

Regular checksum actions should be performed as a preventive measure against bit rot. Given that most universities utilize DSpace, which supports checksums, this can be easily implemented. Checking for bit rot through frequent checksums helps detect data corruption or degradation, which can then trigger appropriate preservation actions to safeguard the long-term integrity of the digital materials.

By addressing these considerations and implementing the suggested actions, IRs can establish robust preservation practices and ensure the long-term viability and accessibility of their digital collections.

### **Monitoring and Evaluation**

Institutional repositories (IRs) should be established with a strong emphasis on implementing a system of ongoing monitoring and evaluation to ensure adherence to established standards and practices. This proactive approach allows for the timely identification and rectification of any gaps or shortcomings that may arise, thus mitigating potential risks or damages.

Furthermore, it is crucial for IRs to aspire to achieve certification as a trusted repository. By seeking such recognition, IRs demonstrate their commitment to maintaining high-quality standards and meeting the expectations set by other reputable digital repositories worldwide. In order to attain this status, comprehensive documentation should be provided, covering all aspects of the repository's operations and procedures. This documentation serves as tangible evidence of the repository's reliability and competence.

Moreover, a formalized process of evaluation within the IR is essential as it guarantees regular and frequent evaluations that ensure the repository remains up-to-date and aligned with evolving best practices in digital preservation and curation. By integrating evaluation as an ongoing practice, IRs can proactively address any potential issues, make necessary improvements, and adapt to changing requirements and technologies.

### **6.5 Recommendations for Further Research**

The aim of this research was to evaluate digital preservation practices in selected institutional repositories within public universities in Kenya. Major gaps were identified from industry best practices. The scope of content covered by the research was limited and therefore further research is recommended in the following areas:

1. The study revealed a lot of gaps in the effectiveness digital preservation practices and therefore there is need to extend the same to born digital information within databases as well as open access journal management systems in Kenyan universities
2. The same kind of research should be extended to the national library being the custodian of Kenya's literary memory in order to accommodate the current trend of digital publishing.
3. The research noted the importance of collaborations in successful digital preservation and therefore would recommend a research to explore its viability in a Kenyan scenario.



## REFERENCES

- Abed, M. G (2015). A consideration to two main ethical issues in educational research, and how may these be addressed. *i-manager's Journal on Educational Psychology*, 8(3), 1-13. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1098173.pdf> on 17/02/2021
- Abrams, S. (2018), Theorizing success: Measures for evaluating digital preservation efficacy. *Presentation at Joint Conference on Digital Library*, Fort Worth. Retrieved from <https://escholarship.org/uc/item/7xt368b2> on 16.02.2021
- Adjei, E., Mensah, M. & Amoafu, E.A. (2019). The story so far-digital preservation in institutional repositories: The case of academic libraries in Ghana. *Digital Library Perspectives*, 35(2), 80-96. DOI 10.1108/DLP-12-2018-0039
- Alex Ball. (2010). *Preservation and Curation in Institutional Repositories* (version 1.3). Edinburgh, UK: Digital Curation Centre.
- Alison, S., McCutcheon, V. & Mahon, M. (2019). Case Study: The University of Glasgow's Digital Preservation Journey 2017-2019. *Insights*, 32 (1): 10. <http://doi.org/10.1629/uksg.461>
- Allinson, J. (2006) OAIS as a reference model for repositories: an evaluation. Report. Retrieved from: <https://core.ac.uk/download/pdf/52154.pdf> on 06/06/2020
- Altman, M. et al (2019). A report on the challenges, opportunities, gaps, emerging trends, and key areas for research and development that support the global capacity for digital stewardship. *2020 NDSA Agenda for Digital Stewardship*. Retrieved from: <https://arxiv.org/ftp/arxiv/papers/2005/2005.05474.pdf> on 03/03/2021
- Anderson, B. (2015). Case Study: University Archives, University of Illinois at Urbana-Champaign. In Marks, S. et al *Becoming a Trusted Digital Repository*. Society of American Archivists p. 60-64. Retrieved from: [https://www2.archivists.org/sites/all/files/Module 8 CaseStudy BethanyAnders on.pdf](https://www2.archivists.org/sites/all/files/Module%208%20CaseStudy%20BethanyAnderson.pdf) on 22/02/2021
- Andriamahady, D. (2021). OAIS compliant digital archiving in DNA. Retrieved from [https://doc.rero.ch/record/333313/files/ANDRIAMAHADY\\_Dina\\_ID\\_OAIS-DNA\\_TB2021.pdf](https://doc.rero.ch/record/333313/files/ANDRIAMAHADY_Dina_ID_OAIS-DNA_TB2021.pdf) on 01/02/2022
- Antunes, G. et al (2012). Assessing Digital Preservation Capabilities Using a Checklist Assessment Method. In *Proceedings of the Ninth International Conference on Preservation of Digital Objects*. Toronto, ON, Canada. Retrieved from: [https://www.researchgate.net/publication/235922156 Assessing Digital Preservation Capabilities Using a Checklist Assessment Method](https://www.researchgate.net/publication/235922156_Assessing_Digital_Preservation_Capabilities_Using_a_Checklist_Assessment_Method) on 22/02/2021
- Anyako, E. N., Echedom, A. U. N., & Baro, E. E. (2018). Digital preservation practices in university libraries: An investigation of institutional repositories in Africa. *Digital Library Perspectives*, 35(1), 41-64. DOI: 10.1108/DLP-10-2017-0041

- ARMA (2015). A Guide to Making Long-Term Digital Preservation Affordable and Sustainable. ARMA.
- Arora, J. (2006). Digital preservation and management: an overview. Paper presented at the 4<sup>th</sup> Convention PLANNER -2006, Mizoram University, Aizawl, INFLIBNET Centre, Ahmedabad.  
assessed from <https://ir.inflibnet.ac.in:8080/ir/bitstream/1944/1187/1/1-20.pdf> on 04/12/2017
- Arora, J. (2009). *Digital Preservation: An Overview*. In Digital Preservation, Management and Access to information in the Twenty First Century (pp. 106-148). Ahmedabad: INFLIBNET Centre. Retrieved from: <https://ir.inflibnet.ac.in/bitstream/1944/1466/1/8.pdf> on 02/06/2019
- Asiamah, N. et al (2017). General, Target, and Accessible Population: Demystifying the Concepts for Effective Sampling. *The Qualitative Report*, 22(6), How to Article 4, 1607-1622. Retrieved from: <https://nsuworks.nova.edu/tqr/vol22/iss6/9on> 13/12/2020
- Austin C.C. et al (2016). Research Data Repositories: Review of Current Features, Gap Analysis, and Recommendations for Minimum Requirements. DOI: <https://doi.org/10.29173/iq904>
- Ayre, C. & Muir, A., (2004). Right to preserve? Copyright and licensing for digital preservation project: final report. Loughborough: Loughborough University. retrieved from: <https://dspace.lboro.ac.uk/2134/343on> 10/12/2020
- Bacciu, C., Angelica L. D. & Andrea M. (2019). The Use of Block chain for Digital Archives: Challenges and Perspectives. AIUCD Annual Conference – Pedagogy, Teaching and Research in the Age of Digital Humanities, 78-81
- Baillieul, J., Grenier, G. & Setti, G. (2018). Reflections on the Future of Research Curation and Research Reproducibility [Point of View] in Proceedings of the IEEE, vol. 106, no. 5, pp. 779-783, doi: 10.1109/JPROC.2018.2816618.
- Ball, A. (2010). Preservation and curation in institutional repositories (version 1.3). Edinburgh, UK: Digital Curation Centre. Retrieved from: [http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/08/recurso\\_02.pdf](http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/08/recurso_02.pdf) on 12/05/2019
- Bantin, P.C. (2016). Resources, Policies and Management structures in Bantan, P.C. (ed) Building trustworthy digital repositories: theory and implementation. Lanham: Rowman & Littlefield
- Barateiro, J. et al (2009). *Addressing Digital Preservation: Proposals for New Perspectives*. In DP'09 First International Workshop on Innovation in Digital Preservation, Texas, USA. Retrieved from: <http://repositorio.lnec.pt:8080/jspui/bitstream/123456789/16408/3/indp2009-7.pdf> on 01/12/19

- Barons, M. et al (2021). Safeguarding the nation's digital memory: towards a Bayesian model of digital preservation risk, *Archives and Records*, 42:1, 58-78, DOI: 10.1080/23257962.2021.1873121
- Barrueco, J.M. & Termens, M. (2022). Digital preservation in institutional repositories: a systematic literature review. *Digital Library Perspectives*, Vol. 38 No. 2, pp. 161-174. doi: 10.1108/DLP-02-2021-0011
- Barve, S. (2007). *File Formats in Digital Preservation*. In A. R. D Prasad & Devika P. Madalli (Eds.), *International Conference on Semantic Web and Digital Libraries* (pp. 239-248). Bangalore: DRTC. Retrieved from: <http://dlissu.pbworks.com/f/File+format1.pdf> on 05/06/2020
- Bashir, M., et al (2008). Reliability and Validity of Qualitative and Operational Research Paradigm. *Pakistan Journal of Statistics and Operation Research (PJSOR)*, 4(1). <https://doi.org/10.18187/pjsor.v4i1.59>
- Baucom, E. (2019) Planning and Implementing a Sustainable Digital Preservation Program. *Library technology reports*, 55(6), 1. Retrieved from: <https://journals.ala.org/index.php/ltr/issue/viewIssue/732/495> on 15/02/2021
- Baucom, E., (2019) A Brief History of Digital Preservation. *Mansfeld Library Faculty Publications*. 31. Retrieved from Retrieved from: [https://scholarworks.umt.edu/ml\\_pubs/31/](https://scholarworks.umt.edu/ml_pubs/31/) on 06/04/2018
- Beagrie, N. et al (2002). Trusted Digital Repositories: Attributes and Responsibilities. Research Libraries Group & Online Computer Library Center, Report. Retrieved from <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf> assessed on 02/11/2019
- Beagrie, N., et al (2008). Digital Preservation Policies Study Part 1: Final Report October 2008, A Study Funded by JISC (Higher Education Funding Council for England (HEFCE): October 2008), 5–13. Retrieved from: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.214.9056&rep=rep1&type=pdf> on 10/12/2020
- Beagrie, N., Rettberg, N. & Williams, P. (2008). *Digital preservation policy: a subject of no importance?* paper presented at iPRES 2008, The Fifth International Conference on Preservation of Digital Objects, 29-30 September 2008, London. Retrieved from: [http://www.bl.uk/ipres2008/presentations\\_day1/01\\_Beagrie.pdf](http://www.bl.uk/ipres2008/presentations_day1/01_Beagrie.pdf) on 07/06/2019
- Becker, C. & Rauber, A. (2011). Decision criteria in digital preservation: what to measure and how. *Journal of the American Society for Information Science and Technology*, 62(6),1009-1028. <https://doi.org/10.1002/asi.21527>
- Becker, C. et al (2009). Systematic planning for digital preservation: Evaluating potential strategies and building preservations plans. *International Journal on Digital Libraries*, 10, 133–157. <https://doi.org/10.1007/s00799-009-0057-1>

- Becker, C., Faria, L. & Duretec, K. (2014). Scalable decision support for digital preservation. *OCLC Systems & Services: International digital library perspectives*, 30(4), 249-284. <https://doi.org/10.1108/OCLC-06-2014-0025>
- Beedham, H. et al. (2005). Assessment of UKDA and TNA compliance with OAIS and METS Standards. Colchester, Essex: UK Data Archive, University of Essex. Retrieved from: [https://iassistquarterly.com/public/pdfs/iqvol36\\_2\\_recker.pdf](https://iassistquarterly.com/public/pdfs/iqvol36_2_recker.pdf) on 30/04/2021
- Bell, E., Bryman, A. & Harley, B., (2018). *Business research methods*. (5th ed). New York: Oxford University Press.
- Berg, H. L. & Lune, B.L. (2017). *Qualitative research methods for social sciences* (9<sup>th</sup>ed). London, Pearson
- Besser, H. (2000). Digital longevity. In Maxine K. Sitts (ed). *Handbook for Digital Projects: a Management Tool for Preservation and Access* (pp 162-176). Andover, Mass.: Northeast Document Conservation Center. Retrieved from: <https://www.nedcc.org/assets/media/documents/dman.pdf> on 03/12/2016
- Bettavia, R. S. (2016). The power of imaginary users: Designated communities in the OAIS reference model. *Proceedings of the Association for Information Science and Technology*, 53(1), 1-9. doi:10.1002/pra2.2016.14505301038
- Bhatacherjee, A. (2012), *Social Science Research: Principles, Methods, and Practices*. Global Text book project
- Blewer, A., Romkey, S., & Spencer, R. (2019). Archivematica as a case study for sustained digital Preservation. *16th International Conference on Digital Preservation iPRES 2019*, Amsterdam, The Netherlands. DOI: 10.1145/nnnnnnn.nnnnnnn
- Borghoff et al. (2006). *Long-Term Preservation of Digital Documents, Principles and Practices*. Springer, London.
- Bote, J & Termens, M. (2019). Reusing Data: Technical and Ethical Challenges. *Journal of Library & Information Technology*, 39(6), 329-337 DOI: 10.14429/djlit.39.6.14807
- Bountouri, L., Gratz, P. & Sanmartin, F. (2018). Digital preservation: how to be trustworthy. *Digital Cultural Heritage*, 10605, 364-374. DOI: 10.1007/978-3-319-75826-8\_29
- Bowen, G.A. (2009), Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 9 (2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- Braun, V., & Clarke, V. (2021). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, 13(2), 201–216. <https://doi.org/10.1080/2159676X.2019.1704846>

- Breeding, M. (May 2002). Digital preservation: building digital collections that will outlast current technologies. *Information Today*.  
<https://librarytechnology.org/document/9718>
- Breytenbach, A., Lourens, A., & Marsh, S. (2013). The role of Jotello F. Soga Library in the digital preservation of South African veterinary history. *Journal of the South African Veterinary Association*, 84(1), 1-7. Retrieved from:  
<http://www.scielo.org.za/pdf/jsava/v84n1/21.pdf> on 08/12/2020
- Brown, A. (2006). The PRONOM PUID Scheme: A Scheme of Persistent Unique Identifiers for Representation Information (Digital Preservation Technical Paper 2). London: The National Archives. Retrieved from  
[https://www.nationalarchives.gov.uk/aboutapps/pronom/pdf/pronom\\_unique\\_identifier\\_scheme.pdf](https://www.nationalarchives.gov.uk/aboutapps/pronom/pdf/pronom_unique_identifier_scheme.pdf) on 02/12/2020
- Brown, A. (2013) *Digital preservation practices: A How-to Guide for Organizations of Any Size*. London. Facet Publishing
- Bryman, A. (2012). *Social science research methods* (4<sup>th</sup> ed.) New York; Oxford University Press
- Budapest Open Access Initiative. (2001). Budapest, Hungary: Open Society Institute. Retrieved from <https://www.budapestopenaccessinitiative.org/read> on 12/09/2018
- Campisi, P., Maiorana, E., Teri, D. & Neri, A. (2009). *Challenges to long-term data preservation: a glimpse of the Italian experience*. In DSP'09: Proceedings of the 16<sup>th</sup> international conference on Digital Signal Processing. Pages 120-127, IEEE Press, Piscataway, NJ. DOI: 10.1109/ICDSP.2009.5201056
- Caplan, P. (2008). Preservation Practices. *Library Technology Reports*, 2, 10-13. Retrieved from: <https://journals.ala.org/index.php/ltr/article/view/4223/4807> on 10/12/2020
- Caplan, P. (2010). The Florida Digital Archive and DAITSS: a model for digital preservation. *Library Hi Tech*, 28(2), 224-234. DOI: 10.1108/07378831011047631
- Caplan, P., & Guenther, R. S. (2005). Practical preservation: the PREMIS experience. *Library Trends*, 54(1), 111-124.  
[https://www.loc.gov/standards/premis/caplan\\_guenther-librarytrends.pdf](https://www.loc.gov/standards/premis/caplan_guenther-librarytrends.pdf)
- Cavaleri, S. A. (2008). Are learning organizations pragmatic? *The Learning Organization*, 15(6), 474-485. <http://dx.doi.org/10.1108/09696470810907383>
- CENDI Digital Preservation Task Group, (2006). Formats for Digital Preservation: A Review of Alternatives and Issues. [http://www.cendi.gov/publications/CENDI\\_PresFormats\\_WhitePaper\\_03092007.pdf](http://www.cendi.gov/publications/CENDI_PresFormats_WhitePaper_03092007.pdf)

- Center for Research Libraries. (2008). Trustworthy repositories audit & certification (TRAC): Criteria and checklist. Chicago, IL: Center for Research Libraries. Retrieved
- Chen S-S. (2001). The Paradox of Digital Preservation. *IEEE Computer*; 34(3), 24–28. <https://doi.org/10.1109/2.910890>
- Chen, Y., & Qu, L. (2008). The Impact of ICT on the Information Economy. 4th International Conference on Wireless Communications, Networking and Mobile Computing, 2008, pp. 1-4, doi: 10.1109/WiCom.2008.2874.
- Chitez, M., Rogobete, R., & Foitoş, A. (2020). Digital Humanities as an Incentive for digitalization Strategies in Eastern European HEIs: A Case Study of Romania. In A. Curaj, L. Deca, & R. Pricopie (Eds.), *European Higher Education Area: Challenges for a New Decade* (pp. 545–564). Springer International Publishing. [https://doi.org/10.1007/978-3-030-56316-5\\_34](https://doi.org/10.1007/978-3-030-56316-5_34)
- Christie, C. A., & Alkin, M. C. (2013). An evaluation theory tree. In M. Alkin (Ed.) *Evaluation roots: A wider perspective of theorists' views and influences*, 2, 11–57.
- CIRL & LoC (2002). *Building a national strategy for digital preservation. Issues in Digital Media Archiving*. Washington, D.C: Council on Library and Information Resources Washington, D.C. and Library of Congress. Retrieved from: <https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub106.pdf> on 10/12/2019
- Cloonan, M., & Sanett, S. (2002). Preservation Strategies for Electronic Records: Where We Are Now—Obliquity and Squint? *The American Archivist*, 65(1), 70-106. Retrieved from <https://www.jstor.org/stable/40294190> on 13/05/2020
- Cohen, L., Manion, L. & Morrison, K. (2007). *Research Methods in Education*. (6<sup>th</sup> ed.). London: Routledge.
- Commission for University Education (2013). Commission for University Education Standards and Guidelines. Retrieved from: <https://cue.or.ke/index.php/downloads1> on 05/04/2018
- Community Archives and heritage Group (2018). Digital Preservation for Community Archives. Retrieved from: <https://www.communityarchives.org.uk/wp-content/uploads/2018/02/Digital-Preservation-for-Community-Archives-V1.4-2018.pdf> on 17/02/2021
- Connaway, L. S., & Powell, R. R. (2010). *Basic research methods for librarians*. (5th ed.). Littleton, CO: Libraries Unlimited.
- Conway, P. (2000). Overview: Rationale for digitization and preservation. In Sitts, M. (Ed.), *Handbook for digital projects: A management tool for preservation and access*. Retrieved from: <http://www.nedcc.org/digital/dighome.htm> On 02/06/2019

- Cornell university library (2010). Digital preservation management long term and short term strategies <http://dpworkshop.org/dpm-eng>
- Corrado, E. M., & Moulaison, H. L. (2014). *Digital preservation for libraries, archives, and museums*. Rowman & Littlefield.
- Corrado, E.M. & Sandy, H.M. (2017) Digital preservation for libraries, archives, and museums. (2<sup>nd</sup>ed). Lanham: Rowman& Littlefield
- Council of Canadian Academies, 2015. *Leading in the DigitalWorld: Opportunities for Canada's Memory Institutions*. Ottawa (ON): The Expert Panel on Memory Institutions and the Digital Revolution, Council of Canadian Academies.[https://cca-reports.ca/wp-content/uploads/2018/10/cofca\\_14-377\\_memoryinstitutions\\_web\\_e.pdf](https://cca-reports.ca/wp-content/uploads/2018/10/cofca_14-377_memoryinstitutions_web_e.pdf)
- Council of Canadians Academies. (2015). Leading in the digital world: Opportunities for Canada's memory institutions. The expert panel on memory institutions in the digital revolution. Ottawa: Council of Canadian Academies.
- Council on Library & Information Resources & the Library of Congress (2002). Building a National Strategy for Preservation: Issues in Digital Media Archiving. Retrieved from <https://www.clir.org/pubs/reports/pub106/contents/#preface> on 17/02/2021
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Sage Publications, Inc
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications, Inc.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles: Sage Publications
- Creswell, J. W.& Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches*. Los Angeles: SAGE
- Creswell, J. W., & Clark, P.V. L. (2011). *Designing and Conducting Mixed Methods Research* (2<sup>nd</sup> ed.). London: Sage Publications Ltd.
- Crow, R. (2002). The case for institutional repositories: a SPARC position paper, retrieved from:<http://hdl.handle.net/10106/24350>on 06/04/2018
- Cunningham, A. (2018). Exploring Digital Preservation in the Cloud. In L. Duranti & C. Rogers (Eds.), *Trusting Records and Data in the Cloud: The Creation, Management, and Preservation of Trustworthy Digital Content* (pp. 179-206). Facet. doi:10.29085/9781783304042.009
- Curdt, C. 2019. Supporting the Interdisciplinary, Long-Term Research Project: Patterns in Soil-Vegetation-Atmosphere-Systems by Data Management Services. *Data Science Journal*, 18(1).5. DOI: <https://doi.org/10.5334/dsj-2019-005>



- da Silva Júnior, L.P. & Borges, M. M. (2017). Digital preservation policies of the institutional repositories at Brazilian Federal Universities. *The Electronic Library*, 35(2), pp. 311-321. <https://doi.org/10.1108/EL-09-2015-0170>
- Dappert A. et al. (Eds). (2016). *Digital Preservation metadata for Practitioners: Implementing PREMIS*. Springer.
- Dar, B.A & Ahmad, S. 2017. Recent trends in digital archiving and preservation: a global perspective. Raleigh: Laxmi book publication.
- Darke, P., et al (1998). Successfully completing case study research: Combining rigor, relevance and pragmatism. *Information Systems Journal*, 8(4), 273- 289. <https://doi.org/10.1046/j.1365-2575.1998.00040.x>
- Dawadi, S., Shrestha, S., & Giri, R. A. (2021). Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms. *Journal of Practical Studies in Education*, 2(2), 25-36 DOI: <https://doi.org/10.46809/jpse.v2i2.20>
- Decman, M. & Vintar, M. (2013). A possible solution for digital preservation of e-government: a centralized repository within a cloud computing framework. *Aslib Proceedings*, 65(4), 5. DOI: 10.1108/AP-05-2012-0049
- Delaney, E. & de Jong, A. (2015). *Media Archives and Digital Preservation: Overcoming Cultural Barriers*, Digital Preservation for the Arts, Social Sciences and Humanities, Dublin, 25-26 June 2015. Dublin: The Digital Repository of Ireland.
- Dell, E. Y. & Shultz, S. M. (2014) Conserving Digital Resources: Issues and Future Access. *Journal of Electronic Resources in Medical Libraries*, 11(3), 124-133, <https://doi.org/10.1080/15424065.2014.937657>
- Deridder J.L. (2016). Digital preservation: Why is this important to me? Retrieved from: <https://www.rjionline.org/stories/digital-preservation-why-is-this-important-to-me> on 15/02/2021
- Deshpande, J. (2016). Digital preservation in library. *International Journal of Library & Information Science (IJLIS)*. 5(2) 25–29. Article ID: IJLIS\_05\_02\_004
- Dharini, A. (2009). *Digital Preservation and Persistence of Digital Collection*. INFLIBNET Center. <http://ir.inflibnet.ac.in/handle/1944/1001>
- Digital Preservation Coalition (2015). Digital Preservation Handbook Technical Solutions and Tools. retrieved from: <https://www.dpconline.org/docs/digital-preservation-handbook2/1550-dp-handbook-technical-solutions-and-tools/file> on 14/12/2020
- Digital Preservation Coalition. (2008). *Preservation Management of Digital Materials: The Handbook*. York: Digital preservation Coalition. Retrieved from <http://www.dpconline.org/advice/preservationhandbook/introduction> on 12/05/2020



- Dobratz, S. et al, (2010). The Use of Quality Management Standards in Trustworthy Digital Archives. *International Journal of Digital Curation* 5(1) DOI: 10.2218/ijdc.v5i1.143
- Dobratz, S., Schoger, A., & Strathmann, S. (2007). The nestor Catalogue of Criteria for Trusted Digital Repository Evaluation and Certification. *Journal of Digital Information*, 8(2). Retrieved from <https://journals.tdl.org/jodi/index.php/jodi/article/view/199/180> on 03/05/2020
- Dollar, C. M. & Ashley, L. J. (2015). Long-Term Digital Preservation. In: R. F. Smallwood *Information Governance: Concepts Strategies and Best Practices*. Wiley. 317-347, <https://doi.org/10.1002/9781118433829.ch17>
- Dollar, C.M & Ashley, L.J (2020) Long term Digital Preservation, in Robert F. Smallwood in *Information Governance: Concepts, Strategies, and Best Practices*, 2<sup>nd</sup> ed, 390-423, John Wiley & Sons, Inc.
- Donaldson, D. R. (2020). Certification Information on Trustworthy Digital Repository Websites: A Content Analysis. *PLoS ONE* 15(12). DOI: 10.1371/journal.pone.0242525
- Dowding, H. (2016). Open Access, Digital Preservation, and the Changing Scholarly Record. In Smith, K. & K.A. Dickson (Eds.), *Open Access and the Future of Academic Libraries*. Lanham, MD: Rowman& Littlefield. Retrieved from: <http://hdl.handle.net/2022/20892> on 10/12/2020
- Downs, R.R. (2019). International Standards for Trustworthy Data Repositories in *National Institutes of Health (NIH) Trustworthy Data Repositories Workshop. National Institute for Allergy and Infectious Diseases (NIAID)*. Rockville, MD, 8–9 Apr 2019. DOI: <https://doi.org/10.7916/d8-h7xf-ha70>
- Dressler, V.A. (2017). The state of affairs with digital preservation at ARL member libraries: A survey and analysis of policy. *Digital Library Perspectives*, 33(2). 137-155. <https://doi.org/10.1108/DLP-08-2016-0030>
- Duff, W.M., Limkilde, C. & van Ballegoie, M. (2006). Digital preservation education: educating or networking? *The American Archivist*. 69, 188-212. DOI: 10.17723/aarc.69.1.0437j178u17665v5
- Dul, J. & Hak, T. (2008) *Case study methodology in business research*, Oxford, Elsevier Ltd.
- Durant, F. (2019). Preservation assessments in Edward, K. A. & Leonard, M. *Analysis and Assessment in Technical Services*. Chicago: ALA Editions retrieved from: <https://ufdc.ufl.edu/IR00011301/00001> on 22/02/2021
- Duranti, L. (2010). The long-term preservation of the digital heritage: A case study of universities' institutional repositories. *JLIS.it*, 1(1), 157–168. Retrieved from: <http://dx.doi.org/10.4403/jlis.it-12>. On 10/12/2019

- El Idrissi, B., (2019). *Long-term digital preservation: A preliminary study on software and format obsolescence*. In Proceedings of the ArabWIC 6th Annual International Conference Research Track (pp. 1-6). March, 2019. (Conference Proceedings)
- Electronic Information for Libraries (EIFL). (2018). How EIFL's support helped open up East African research to the world: Seven years of EIFL's work in Kenya, Tanzania and Uganda. Retrieved from: <https://www.eifl.net/resources/how-eifl-support-helped-open-east-african-research-world> on 14/11/2019
- Emmott, S. (2008). Preservation of Web Resources: Making a Start. *ARIADNE*. 56. Retrieved from: <http://www.ariadne.ac.uk/issue/56/jisc-powr-rpt/> on 11/12/2020
- Engelhardt, C (2013). *The DigCurV review of training needs in the field of digital preservation and curation*. An overview of the main findings. In: CirinnàCh, Fernie K, Lunghi M, Casarosa V (eds) Proceedings of the framing the digital curation curriculum conference. CEUR workshop proceedings, vol 1016. CEUR-WS, Aachen. Retrieved from: <http://ceur-ws.org/Vol-1016/paper10.pdf> on 10/12/2020
- Erima, J., Masai, W., & Wosyanju, M. G. (2016). *Preservation of digital research content in academic institutions: A case study of Moi University, Kenya*. 2016 IST-Africa Week Conference, 1–11. <https://doi.org/10.1109/ISTAFRICA.2016.7530620>
- Fauduet, L & Peyrard, S. (2010) *A data-first preservation strategy: data management in spar*. In IPres 2010 Viena Austria. <https://phaidra.univie.ac.at/o:185415>
- Faundeen, J. (2017) Developing Criteria to Establish Trusted Digital Repositories. *Data Science Journal*, 16 (22), 1–13, <https://doi.org/10.5334/dsj-2017-022>
- Feeney, M. (1999). Towards a National Strategy for Archiving Digital Materials. *Alexandria*, 11(2), 107–122. <https://doi.org/10.1177/095574909901100204>
- Flathers, E., Kenyon, J. & Gessler, P.E. (2017). A service-based framework for the OAIS model for earth science data management. *Earth Sci Inform* 10, 383–393. <https://doi.org/10.1007/s12145-017-0297-3>
- Fleischhauer, C. (2000). Archival Information Package (AIP) Design Study. retrieved from: [https://www.loc.gov/rr/mopic/avprot/AIP-Study\\_v19.pdf](https://www.loc.gov/rr/mopic/avprot/AIP-Study_v19.pdf) on 06/07/2020
- Florida Center for Library Automation (FLCA), (2011). DAITSS Digital Preservation Repository Software. Florida Center for Library Automation (FCLA). Accessed from: <https://daitss.fcla.edu/> on 02/06/2019
- Fornaro, P., Rosenthaler, L. & Zbinden, E (2017). TIFF in Archives: A survey about existing files in memory institutions. *IS&T Archiving*, pp 6 - 10, <https://doi.org/10.2352/issn.2168-3204.2017.1.0.6>

- Francke, H., Gamalielsson, J. & Lundell, B. (2017). Institutional repositories as infrastructures for long-term preservation. *Information Research*, 22(2), paper 757. Retrieved from: <http://InformationR.net/ir/22-2/paper757.html> on 10/12/2020
- Frank, R. D. (2018). The Social Construction of Risk in Trustworthy Digital Repository Audit and Certification. [unpublished Masters Thesis], University of Michigan. Retrieved from <https://deepblue.lib.umich.edu/handle/2027.42/147539> on 14/12/2020
- Friese, Y. (2012), How to develop a preservation policy. *International Association for Social Science Information Service and Technology (IASIST) Quarterly*, 36(3/4), 17-22. Retrieved from: [https://iassistquarterly.com/public/pdfs/iqvol36\\_34\\_friese.pdf](https://iassistquarterly.com/public/pdfs/iqvol36_34_friese.pdf) on 10/12/2020
- Gallinger, M. et al (2017). Trends in Digital Preservation Capacity and Practice: Results from the 2nd Bi-annual National Digital Stewardship Alliance Storage Survey. *D-Lib Magazine*, 23(7/8). <https://doi.org/10.1045/july2017-gallinger>
- Gartner, R. & Lavoie, B., (2013). Preservation Metadata (2nd edition). *DPC Technology Watch Report 13-3 May 2013*. <http://dx.doi.org/10.7207/twr13-03>
- Gaur, R., & Tripathi, M. (2012). Digital Preservation of Electronic Resources. *DESIDOC Journal of Library & Information Technology*, 32(4). <https://doi.org/10.14429/djlit.32.4.2522>
- Gbaje E.S. (2011), Digital preservation strategies: A case study of Nigerian national information centres. *IFLA Journal*, 37(3):218-227. <https://doi.org/10.1177/0340035211418727>
- Gbaje, E. S., & Mohammed, M. F. (2017). Long-term accessibility and re-use of institutional repository contents of some selected academic institutions in Nigeria. <https://library.ifla.org/id/eprint/1943/1/S12-2017-gbaje-en.pdf>
- Genoni, P. (2004). Content in institutional repositories: a collection management issue. *Library Management*, 25(6-7), 300-306. <https://doi.org/10.1108/01435120410547968>
- Gerring, J. (2007). *Case Study Research: Principles and Practices*. (2nd ed.). New York. Cambridge University Press.
- Giaretta, D. (2011). *Advanced digital preservation*. London: Springer. DOI 10.1007/978-3-642-16809-3
- Godfrey-Smith, P. (2002). Dewey on naturalism, realism and science. *Philosophy of Science*, 69(S3), S25–S35. doi: 10.1086/341765
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *Qualitative Report*, 8, 597-606. Retrieved from: <https://core.ac.uk/download/pdf/51087041.pdf> on 06/12/2020

- Goldkuhl, G. (2004). Meanings of pragmatism: Ways to conduct information systems research. Action in Language. *Organisations and Information Systems*. 13-26. Retrieved from: [https://www.researchgate.net/profile/Goeran-Goldkuhl/publication/246789082\\_Meanings\\_of\\_Pragmatism\\_Ways\\_to\\_conduct\\_information\\_systems\\_research/links/5592907608ae16f493ee17a7/Meanings-of-Pragmatism-Ways-to-conduct-information-systems-research.pdf](https://www.researchgate.net/profile/Goeran-Goldkuhl/publication/246789082_Meanings_of_Pragmatism_Ways_to_conduct_information_systems_research/links/5592907608ae16f493ee17a7/Meanings-of-Pragmatism-Ways-to-conduct-information-systems-research.pdf) on 10/10/2022
- Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*. 21: 135–46. DOI: <https://doi.org/10.1057/ejis.2011.54>
- Göldner, R. (2017). Archiving Archaeological Research Data. On Requirements, Objectives and First Experiences from a German Respectively Saxon Point of View. *SDH*, 1, 2, 165-181. DOI:10.14434/sdh.v1i2.23205
- Graneheim, U. H., & Lundman, B. (2004). Qualitative Content analysis in nursing research: Concepts, procedures, and measures to achieve trustworthiness. *Nurse Education Today* 24(2), 105-112.
- Gross, J.M.S (2018). Document Analysis. In: Bruce B. Frey ed. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. New Delhi; Sage publications. Pp 1406-1414
- Guest G, Namey E, Chen M (2020) A simple method to assess and report thematic saturation in qualitative research. *PLoS ONE*,15(5): e0232076. <https://doi.org/10.1371/journal.pone.0232076>
- Hammersley, M. & Traianou, A. (2012). What is ethics? In *Ethics in qualitative research: Controversies and contexts* (pp. 16-34). SAGE Publications Ltd, <https://www.doi.org/10.4135/9781473957619>
- Hardin S. (2008). Plenary Address Report: The Role of Universities in the Preservation of Cultural Memory. *Bulletin of the American Society for Information Science and Technology*, 34(4). Available at: <https://asistdl.onlinelibrary.wiley.com/doi/pdfdirect/10.1002/bult.2008.1720340407> on 12/12/2020
- Harvey, D. (2005). *Preserving Digital Materials*. München: KG Saur.
- Harvey, R. & Mahard, M. (2013). Mapping the Preservation Landscape for the Twenty-First Century. *Preservation, Digital Technology & Culture*, 42(1), 5–16. <https://doi.org/10.1515/pdte-2013-0002>
- Harvey, R., (2011). *Preserving Digital Materials*, (2<sup>nd</sup> ed.). De Gruyter Saur.
- Hedstrom, M. & Montgomery, S. (1998), Digital preservation needs and requirements in RLG member institutions. *Research Library Group*, Mountain View, CA, working paper, December. Retrieved from: <https://www.oclc.org/content/dam/research/activities/digpresneeds/digpres.pdf> on 11/09/2018

- Hedstrom, M. (1998). Digital Preservation: A Time Bomb for Digital Libraries. *Computers and the Humanities* 31 189–202 Retrieved from: <https://www.jstor.org/stable/30200423> on 12/05/2020
- Hedstrom, M. (2001). Digital preservation: problems and prospects. *Digital Libraries*, 20. retrieved from: <https://core.ac.uk/download/pdf/56647731.pdf> on 10/12/2020
- Hedstrom, M. et al (2003). Invest to Save: Report and Recommendations of the NSF-DELOS Working Group on Digital Archiving and Preservation. DOI: 10.13140/RG.2.1.2813.6089
- Hennink, M., Hutter, I., & Bailey, A. (2020). *Qualitative Research Methods*. SAGE Publications
- Hitchcock, S., et al (2005) Preservation for Institutional Repositories: practical and invisible. Ensuring Long-term Preservation and Adding Value to Scientific and Technical data (PV 2005), *The Royal Society of Edinburgh*, Edinburgh, Scotland, 2005-11-21 - 2005-11-23 <http://eprints.soton.ac.uk/id/eprint/18774>
- Hockx-Yu, H. & Brower, D. (2018) Typology of Digital Collection as a Framework for Digital Preservation. Short paper. *16th International Conference on Digital Preservation iPRES*, Amsterdam, The Netherlands. DOI 10.17605/OSF.IO/BPXMV
- Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. *Program: electronic library and information systems*, 40(3), 232-243. <https://doi.org/10.1108/00330330610681312>
- Hoeren, T. et al (2013). *Legal aspects of digital preservation*. Cheltenham, UK. Edward Elgar Publishing.
- Hoke, G. E. (2012). Future Watch: Strategies for Long-Term Preservation of Electronic Records. *Information Management*, 1-7. Retrieved from: <https://www.dar.cah.ucf.edu/items/show/66> on 03/06/2020
- Howell, A. (2000). Perfect One Day—Digital The Next: Challenges in Preserving Digital Information. *Australian Academic & Research Libraries*, 31(4), 121-141. DOI: 10.1080/00048623.2000.10755130
- Hsieh, H & Shannon, S. E. (2005) Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*. 15(9), 1277-1288. DOI: 10.1177/1049732305276687
- Hurley, G., & Shearer, K. (2019). Final Report of the Survey on Digital Preservation Capacity and Needs at Canadian Memory Institutions, 2017-18. Canadian association of research libraries. doi: 10.5683/SP2/8KLEPP
- International Records Management Trust (2016). Understanding Digital Records Preservation Initiatives. The international council of archives. Retrieved from: [https://www.ica.org/sites/default/files/Digital%20Preservation%20Initiatives%20Module\\_0.pdf](https://www.ica.org/sites/default/files/Digital%20Preservation%20Initiatives%20Module_0.pdf) on 11/06/2019

- International Standard Organisation- ISO 16363:2012: *Space data and information transfer systems — Audit and certification of trustworthy digital repositories* retrieved from:  
<https://www.kdrs.no/prosjekt/ProsessDigitaltDepot/V2.02/Dokumenter/ISO16363.pdf> on 04/03/2017
- InterPARES Project. (2008). The International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 3 Project. retrieved from: <http://www.interpares.org> on 18/06/2019
- Interuniversity Consortium for Political and Social Research (ICPSR) (2009). *Principles and Good Practice for Preserving Data*. International Household Survey Network, IHSN Working Paper No 003. Retrieved from: <http://ihsn.org/sites/default/files/resources/IHSN-WP003.pdf> on 10/12/2020
- Ismail, A & Affandy, H. B. (2018). *Conceptual Paper: Digital Preservation Strategies in Archival Institution*. MATEC Web of Conferences 150, 05052 (2018). <https://doi.org/10.1051/mateconf/201815005052>
- ISO 14721:2012: Space data and information transfer systems — Open archival information system (OAIS) — Reference model. 2nd ed. International standard organisation ISO.
- ISO 16363:2012: Space data and information transfer systems — Audit and certification of trustworthy digital repositories. International standards Organisation. Retrieved from: <https://www.kdrs.no/prosjekt/ProsessDigitaltDepot/V2.02/Dokumenter/ISO16363.pdf> on 10/12/2020
- Jain, P. (2012). Promoting open access to research in academic libraries. *Library Philosophy and Practice*, 737(1), 1-13. Retrieved from: <https://digitalcommons.unl.edu/libphilprac/737/> on 06/04/2018
- Jharota, A.K. (2018). *Importance of Digital Preservation in Digital Era*. paper presented in National Conference on Library Information Science & Information Technology for Education held on 27 August 2017, organized by MRES, Delhi, pp.12-15. Retrieved from: [https://www.researchgate.net/publication/326534604\\_IMPORTANCE\\_OF\\_DIGITAL\\_PRESERVATION\\_IN\\_DIGITAL\\_ERA](https://www.researchgate.net/publication/326534604_IMPORTANCE_OF_DIGITAL_PRESERVATION_IN_DIGITAL_ERA) on 16/02/2021
- Johnson, R. (2002), Institutional repositories: partnering with faculty to enhance scholarly communication, *D-Lib Magazine*, Vol. 8 No. 11, Retrieved from: <http://www.dlib.org/dlib/november02/johnson/11johnson.html> on 14/05/2019
- Johnstone, L. (2020). Challenges in preservation and archiving digital materials. L Johnston. *Information Services & Use*. 40 (3), 193-199 DOI: 10.3233/ISU-200090



- Jomo Kenyatta University Agriculture and Technology. (2015). digital repository policy, retrieved from: <http://jkuat.ac.ke/departments/library/wp-content/uploads/2015/04/POLICY-JKUAT-Digital-Repository-Policy31.pdf> on 12/08/2017
- Jomo Kenyatta University of Science and Technology (2022). JKUAT repository. Accessed from: <http://ir.jkuat.ac.ke/>. on 10/10/2022
- Jones, M. & Semple, N. (2006). Mind the Gap: Digital Preservation Needs in the UK. *Ariadne*, 48. Retrieved from: <http://www.ariadne.ac.uk/issue/48/semple-jones> on 06/06/2019
- Jones, M. (2006). The Digital Preservation Coalition: Building a National Infrastructure for Preserving Digital Resources in the UK. *The Serials Librarian*, 49(3), 95-104. [https://doi.org/10.1300/J123v49n03\\_08](https://doi.org/10.1300/J123v49n03_08)
- Joo, S., Hofman, D. & Kim, Y. (2019). Investigation of challenges in academic institutional repositories: A survey of academic librarians. *Library Hi Tech*, 37(3), 525-548. DOI:10.1108/lht-12-2017-0266
- Kabir, S. M. S. (2016). *Basic guidelines for research: An Introductory Approach for All Disciplines*. Chittagong – 4203: Book Zone Publication
- Kalusopa, T (2018). Preservation and Access to Digital Materials: Strategic Policy Options for Africa. DOI: 10.4018/978-1-5225-6921-3.ch021 150-174
- Kaplan, B. & Maxwell, J. (2005). Qualitative Research Methods for Evaluating Computer Information Systems, in *Evaluating the Organizational Impact of Healthcare Information Systems*. New York; Springer: New York. pp. 30-55.
- Kastellec, M. (2012). Practical Limits to the Scope of Digital Preservation. *Information Technology and Libraries* 31:63–71. DOI: <https://doi.org/10.6017/ital.v31i2.2167>
- Kavishe, G.F & Dulle, F. (2016). Preservation skills and strategies of electronic information resources in the digital era: Case of University of Kwazulu-Natal Libraries. *Library Philosophy and Practice*. Retrieved from: <https://core.ac.uk/download/pdf/77940903.pdf> on 19/02/2021
- Keller, S.R. et al (2019) Digital Preservation Task Force Update. *The Serials Librarian*, 76(1-4), 51-54. <https://doi.org/10.1080/0361526X.2019.1583545>
- Kelly, L.M. & Cordeiro, C. (2020). Three principles of pragmatism for research on organizational processes. *Methodological Innovations*. May-August 2020: 1–10 <https://doi.org/10.1177/2059799120937242>
- Kemenade, J. V. (2020), *The CERN Digital Memory Platform: Building a CERN scale OAI compliant Archival Service* (Masters thesis). Vrije University Amsterdam <https://cds.cern.ch/record/2728246/files/CERN-THESIS-2020-092.pdf>
- Kenyatta University (2022). Kenyatta University repository home. available at: <https://ir-library.ku.ac.ke/>. Accessed on 01/10/2022

- Kenyatta University. (2013). Open access institutional repository policy, retrieved from: <http://library.ku.ac.ke/wp-content/uploads/2013/01/Library-IR-Policy.pdf> on 12/08/2017
- Kibiswa, N. K. (2019). Directed Qualitative Content Analysis (DQICA): A Tool for Conflict Analysis. *The Qualitative Report*, 24(8), 2059-2079. <https://doi.org/10.46743/2160-3715/2019.3778>
- Kilbride, W. G. & Norris, S.L. (2013). Making Our Digital Memory Accessible Tomorrow by Watching Technology Today, *New Review of Information Networking*, 18:2, 70-73, DOI: 10.1080/13614576.2013.839393
- Kirchoff, A. J. (2008). Digital preservation: challenges and implementation. *Learned Publishing*, 21(4), 285-294. <https://doi.org/10.1087/095315108X356716>
- Kivunja, C. & Kuyini, A.B. (2017). Understanding and Applying Research Paradigms in Educational Contexts. *International Journal of Higher Education*, 6(5). 26-41. doi:10.5430/ijhe.v6n5p26
- Knight, G. (2005). *SHERPA DP: establishing an OAIS-compliant preservation environment for institutional repositories*. In *Digital repositories: interoperability and common services*, 9th DELOS thematic workshop, Heraklion, Crete, May 11- 13, 2005 (pp. 43-48). Retrieved from: <http://delos-wp5.ukoln.ac.uk/dissemination/pdfs/knight.pdf> on 02/06/2019
- Kolle, R.S., Reddy, P.B., Parmeshwarand, S & Basavaraj, S. (2014). Strategies and Techniques for Preservation of Digital Resources. *Pearl: A Journal of Library and Information Science*, 8(4), 221- 225. DOI: 10.5958/0975-6922.2014.00747.5
- Kool, W., Van der Werf, T. & Lavoie, B. (2014). *Preservation Health Check: Monitoring Threats to Digital Repository Content*. Dublin, Ohio: OCLC Research. Retrieved from <http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch/preservation-health-check-2014.pdf> on 01/12/2020
- Korb, J. & Strodl, S. (2010). Digital preservation for enterprise content: a gap-analysis between ECM and OAIS. In: *proceedings of the 7th International Conference on Preservation of Digital Objects (iPRES 2010)*, Wien, pp. 221–229
- Kremser, J. et al (2012). Long Term Preservation of Digital Data: Background Research, CESNET Technical Report 2/2012. <https://www.cesnet.cz/wp-content/uploads/2013/01/digital-data-long-term-preservation.pdf>
- Labrooy M. (2013) Open Access Nairobi. E learning Africa News: Perspectives ICT, education and Development in Africa. available at: [http://www.elearning-africa.com/eLA\\_Newsportal/open-access-nairobi/](http://www.elearning-africa.com/eLA_Newsportal/open-access-nairobi/)
- Lakshmi, V. & Jindal, S. C. (eds.). (2004). *Digital libraries*. Delhi: Isha Books.



- Lampert, C. & Vaughan, J. (2018). Preparing to Preserve: Three Essential Steps to Building Experience with Long-Term Digital Preservation. IFLA WLIC 2018 Kuala Lumpur, Malaysia: retrieved from: [https://digitalscholarship.unlv.edu/lib\\_articles/669](https://digitalscholarship.unlv.edu/lib_articles/669) on 04/04/2021
- Langley, S. (2019). Digital Preservation Should Be More Holistic: A Digital Stewardship Approach. In Myntti, J. and Zoom, J. (eds) *Digital Preservation in Libraries: Preparing for a Sustainable Future*. Chicago: American Library Association. Retrieved from: <https://www.repository.cam.ac.uk/bitstream/handle/1810/287006/EDITIONS%20ALCTS%20Digital%20%20CH%207.pdf?sequence=1&isAllowed=yon> 05/04/2020
- Lavoie, B. (2004). The Open Archival Information System Reference Model: Introductory Guide. *De Gruyter*, 33(2), 68-81. <https://doi.org/10.1515/MFIR.2004.68>
- Lee, C. A. (2005). Defining Digital Preservation Work: A Case Study of the Development of the Reference Model for an Open Archival Information System. [Unpublished PhD Dissertation], University of Michigan. Retrieved from: [https://ils.unc.edu/caltee/dissertation\\_caltee.pdf](https://ils.unc.edu/caltee/dissertation_caltee.pdf) on 05/12/2020
- Lee, C. A. (2010). Open Archival Information System (OAIS) Reference Model. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences*. (3rd Ed., pp. 4020–4030). Taylor & Francis. doi:10.1081/E-ELIS3-120044377
- Lee, C., & Tibbo, H. (2007). Digital Curation and Trusted Repositories: Steps Toward Success. *Journal of Digital Information*, 8(2). Retrieved from <https://journals.tdl.org/jodi/index.php/jodi/article/view/229/183> on 13/11/2019
- Lee, D.J. & Stvilia, B. (2017) Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PLoS ONE* 12(3), 1-44. <https://doi.org/10.1371/journal>.
- Lee, K. H., et al (2002). The State of the Art and Practice in Digital Preservation. *Journal of research of the National Institute of Standards and Technology*, 107(1), 93–106. doi:10.6028/jres.107.010
- Leedy, P. D. et al (2019). *Practical Research: Planning and Design*, (12<sup>th</sup> ed.) London, Pearson
- Library of congress (2010) Preserving Our Digital Heritage: The National Digital Information Infrastructure and Preservation Program 2010 Report. [https://www.digitalpreservation.gov/multimedia/documents/NDIIPP2010Report\\_Post.pdf](https://www.digitalpreservation.gov/multimedia/documents/NDIIPP2010Report_Post.pdf)
- Lin, D., et al (2020). The TRUST Principles for digital repositories. *Scientific data*, 7(1), 144. <https://doi.org/10.1038/s41597-020-0486-7>

- Lorie, R.A. (2001). Long term preservation of digital information. JCDL '01: Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries Roanoke. January 2001 Pages 346–352 <https://doi.org/10.1145/379437.379726>
- Luan, F., Nygard, M. & Mestl, T. (2010). A survey of digital preservation strategies. *World Digital Libraries-An international journal*, 3(2), 133–150. Retrieved from: <https://research.idi.ntnu.no/longrec/papers/WDL.pdf> on 12/06/2019
- Lundell, B. (2012). *Why do we need Open Standards?* In Proceedings: EURAS Annual Standardization Conference ‘Standards and Innovation’ (pp. 227–240). Aachen: G. Mainz Verlag. Retrieved from: <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A551564&dswid=-4515> on 10/12/2020
- Lunghi, M., Grindley, N., Stoklasova, S., Trehub, A. & Egger, C. (2012). *Economic Alignment*. In McGovern, N. Y’s Aligning National Approaches to Digital Preservation, (195–234). Atlanta, GA: Educopia Institute Publications. Retrieved from: <https://aurora.auburn.edu/bitstream/handle/11200/48584/ANADP%202011%20Economic%20Alignment.pdf?sequence=1&isAllowed=y> on 30/11/2020
- Lynch, C. (1999). Canonicalization: A Fundamental Tool to Facilitate Preservation and Management of Digital Information. *D-Lib Magazine*, 5(9),73–84. Retrieved from: <http://www.dlib.org/dlib/september99/09lynch.html> on 03/12/2019
- Lynch, C. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL Bimonthly Report* 226. <http://www.arl.org/newsltr/226/ir.html>
- Mackenzie, N. & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in educational research*, 16(2), 193-205
- Madill, A. (2008). Realism in Lisa M. Given (ed). In Given, G,(ed), *The Sage Encyclopedia of Qualitative Research Methods*. Thousand Oak, SAGE Publications, Inc. 732-736. DOI: <http://dx.doi.org/10.4135/9781412963909.n370>
- Madsen, C. M. & Hurst, M. (2019). Digital Preservation Policy and Strategy: Where Do I Start? In J. Myntti, & J. Zoom (Eds.), *Digital Preservation: Preparing for a Sustainable Future* (pp. 37–48). ALCTS Monographs. [https://www.researchgate.net/publication/340021081\\_Digital\\_Preservation\\_Policy\\_And\\_Strategy\\_Where\\_Do\\_I\\_Start\\_in\\_Digital\\_Preservation\\_in\\_Libraries\\_Preparing\\_for\\_a\\_Sustainable\\_Future\\_Edited\\_by\\_Jeremy\\_Myntti\\_and\\_Jessalyn\\_Zoom](https://www.researchgate.net/publication/340021081_Digital_Preservation_Policy_And_Strategy_Where_Do_I_Start_in_Digital_Preservation_in_Libraries_Preparing_for_a_Sustainable_Future_Edited_by_Jeremy_Myntti_and_Jessalyn_Zoom)
- Maemura, E. et al (2017). Organizational assessment frameworks for digital preservation: A literature review and mapping. *Journal of the Association for Information Science and Technology (JASIST)*, 68(7), 1619-1637. DOI <https://doi.org/10.1002/asi.23807>

- Magama, B. (2017). Strategies for preservation of digital records in Masvingo Province of Zimbabwe. [unpublished Masters Thesis]. University of South Africa. Retrieved from [http://uir.unisa.ac.za/bitstream/handle/10500/24379/dissertation\\_magama\\_b.pdf?sequence=1](http://uir.unisa.ac.za/bitstream/handle/10500/24379/dissertation_magama_b.pdf?sequence=1) on 22/02/2021
- Maniatis, P. (2005). The LOCKSS peer-to-peer digital preservation system. *ACM Transactions on Computer Systems*, 23(1):2-50. DOI: 10.1145/1047915.1047917
- Margot, N. (2019) Digital Archives: Choosing Sustainable File Formats. Retrieved from <https://www.margotnote.com/blog/2019/06/03/choosing-sustainable-file-formats> on 09/12/2020
- Masenya, T.M. & Ngulube, P. (2019). Digital preservation practices in academic libraries in South Africa in the wake of the digital revolution. *South African Journal of Information Management* 21(1), a1011. <https://doi.org/10.4102/sajim.v21i1.1011>
- Mauthner, N. (2013). Open access data sharing policies: Implications for academic roles, practices and identities. *Society for Research into Higher Education*. Retrieved from: <http://www.srhe.ac.uk/downloads/MauthnerScopingReport.pdf> on 10/05/2020
- Maxwell J. A. (2012). What is realism, and why should qualitative researchers care? In: Maxwell JA, (ed). *A realist approach for qualitative research*. Thousand Oaks, CA: Sage. P. 3–13.
- Mayring, P. (2014) Qualitative content analysis: theoretical foundation, basic procedures and software solution. *Klagenfurt: Monograph*. Retrieved from: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173> on 15/02/2021
- McGovern, N. Y., & McKay, A. C. (2008). Leveraging Short-term Opportunities to Address Long-term Obligations: A Perspective on Institutional Repositories and Digital Preservation Programs. *Library Trends*, 57(2), 262–279. <https://doi.org/10.1353/lib.0.0038>
- McGregor, S. L. (2018). Philosophical underpinnings of the transdisciplinary research methodology. *Transdisciplinary Journal of Engineering & Science*, 9. DOI: <https://doi.org/10.22545/2018/00109>
- McGuirk, K. (2016) AA06: Trusting Digital Archiving Services: Literature Review. Interparestrust. Retrieved from [http://interparestrust.org/assets/public/dissemination/AA06\\_20160821\\_Trusting\\_DigitalArchivingServices\\_LitReview\\_v4.pdf](http://interparestrust.org/assets/public/dissemination/AA06_20160821_Trusting_DigitalArchivingServices_LitReview_v4.pdf) on 10/06/2021
- Mertens, D.M. & Wilson, A.T. (2012), *Program Evaluation Theory and Practice*. New York, Guilford.

- Micunovic, et al (2016). Data organization and preservation in the context of digital and networked media: public's attitudes, habits and practices in relation to digital curation of personal digital data. *Libellarium: Journal for the history of writing, books and memory institutions*. 9(2), 0-0. Retrieved from <https://hrcaak.srce.hr/176677> on 04/03/2021
- Milam, Z. T. (2014). Digital preservation - Requirements for preserving authentic, reliable and usable records in an EDRMS over the long term: A case study. [Unpublished Master Thesis] University of Dundee. DOI: 10.13140/RG.2.2.32473.62568
- Moffat S. (2015). Contextualizing Scientific Research Methodologies. *OSR Journal of Research & Method in Education (IOSR-JRME)*, 5(6), PP 52-57, DOI: 10.9790/7388-05615257
- Mohanty, R., & Das, R. K. (2014). *Digital libraries: Reshaping Traditional Libraries into Next Generation Libraries*. Mumbai: Allied Publishers. Accessed from: <https://books.google.co.ke/books?id=XluhCgAAQBAJ> on 12/12/2020
- Morgan D (2014) Pragmatism as a paradigm for social research. *Qualitative Inquiry*, 20(8): 1045–1053.
- Morgan, S. J. (2017). Case Study Observational Research: A Framework for Conducting Case Study Research Where Observation Data Are the Focus. *Qualitative Health Research*. 27(7) 1060–1068. DOI: 10.1177/1049732316649160
- Moseti, I. (2016). Digital preservation and institutional repositories: case study of universities in Kenya. *Journal of the South African Society of Archivists*, 49, 137-154. Retrieved from: <https://www.ajol.info/index.php/jsasa/article/view/138434> on 05/06/2020
- Mukherjee, S. & Das, R. (2020). Integration of Domain-Specific Metadata Schema for Cultural Heritage Resources to DSpace: A Prototype Design. *Journal of Library Metadata*, 20(2–3),155–78. DOI: 10.1080/19386389.2020.1834093.
- Murgu, c. (2021). Web Archives, Critical Digital Literacy, and the Growing Primacy of Born Digital Objects in Elizabeth Brookbank & Jess Haigh (eds) *Critical Library Pedagogy in Practice* 263-283 assessed from: <https://dr.library.brocku.ca/bitstream/handle/10464/15292/Murgu%20-%20web%20archives.pdf?sequence=1>
- Myntti, J. (2019). *Digital Preservation in Libraries: Preparing for a Sustainable Future* (An ALCTS Monograph). American Library Association.
- Nadal, J. (2007) The Human Element in Digital Preservation. *Collection Management*, 32:3-4, 289-303, DOI: [https://doi.org/10.1300/J105v32n03\\_04](https://doi.org/10.1300/J105v32n03_04)
- Najar J.K. & Wani, J.A (2019). Digital Preservation: An Overview. *Library Philosophy and Practice* <https://digitalcommons.unl.edu/libphilprac/2989>

- National Information Standards Organization (NISO), (2010). Special Issue: Digital Preservation. *Information Standards Quarterly*, 22(2). Retrieved from: <https://www.niso.org/sites/default/files/2019-07/ISQ%20Spring%202010.pdf> on 05/04/2019
- National Information Standards Organization. (2007). A framework of guidance for building good digital collections. (3rd ed.). Bethesda, MD: NISO. Retrieved from: <https://www.niso.org/sites/default/files/2017-08/framework3.pdf> on 06/07/2019
- Nguyen, Q. L., & Lake, A. (2011). *Content server system architecture for providing differentiated levels of service in a digital preservation cloud*. 2011 IEEE International Conference on Cloud Computing (CLOUD): 557-564
- Niehof, J. M. et al (2018), Long-term Preservation of Deprecated Media: How Can Libraries Provide Information from Today's CD-ROMs in the Future? *Paper presented at 2018 ASEE Annual Conference & Exposition*, Salt Lake City, Utah. Retrieved from: <https://peer.asee.org/30777> on 04/03/2021
- Noonan, D. (2014). Digital Preservation Policy Framework: A Case Study. *EDUCAUSE Review* 1–23. Retrieved from: <http://www.educause.edu/ero/article/digital-preservation-policy-framework-case-study> on 11/11/2020 .
- Northeast Document Conservation Center (2019). Digital preservation assessment handbook. Andover, Massachusetts. Retrieved from: <https://www.nedcc.org/assets/media/documents/nedcc-DPA-hndbk-6.24-web.pdf> on 01/03/2021
- Onwuegbuzie, A. J., & Collins, K. M. (2007). A Typology of Mixed Methods Sampling Designs in Social Science Research. *The Qualitative Report*, 12(2), 281-316. Retrieved from <https://nsuworks.nova.edu/tqr/vol12/iss2/9> on 02/12/2020
- OpenDOAR: directory of open access repositories. Accessed at: [http://www.open\\_doar.org/](http://www.open_doar.org/) on 01/02/2022
- Osang, J. E. Et al (2013). Methods of Gathering Data for Research Purpose and Applications Using IJSER Acceptance Rate of Monthly Paper Publication (March 2012 Edition-May 2013 Edition). *IOSR Journal of Computer Engineering (IOSR-JCE)*, 15(2), 59-65. Retrieved from: <http://www.iosrjournals.org/iosr-jce/papers/Vol15-issue2/I01525965.pdf?id=7568> on 12/12/2020
- Owens, T. (2018), *The Theory and Craft of Digital Preservation*. Baltimore, MD: Johns Hopkins University Press
- Palinkas, L. A. et al (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and policy in mental health*, 42(5), 533–544. <https://doi.org/10.1007/s10488-013-0528-y>

- Patel, H. (2014). Backup practices in digital libraries of research and higher learning institutes in Gujarat. [Unpublished Doctoral Thesis]. Sardar Patel University. Retrieved from <http://hdl.handle.net/10603/34674> on 01/06/2020
- Patton M (2005), *Qualitative Research & Evaluation Methods* (4th ed.). Los Angeles, CA: SAGE.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods*. (3rd ed.). Saint Paul, MN: Sage Publications
- Patton, M. Q. (2014). *Qualitative research & evaluation methods*. (4<sup>th</sup> ed.). Saint Paul, MN: Sage Publications
- Pearson, D. & Webb, C. (2008). Defining File Format Obsolescence: A Risky Journey. *The International Journal of Digital Curation*. 1(3), 89-106. DOI: <https://doi.org/10.2218/ijdc.v3i1.44>
- Pearson, D., & Coufal, L. (2013). Digital preservation environment matrix. National & State Libraries Australasia. Retrieved from: <https://www.nsla.org.au/sites/default/files/documents/nsla.digpres-environment-maturity-matrix.pdf> on 21/02/2021
- Pennock, M. (2006) for Joint Information Systems Committee (JISC), Digital Preservation: Continued Access to Authentic Digital Assets accessed from [http://www.jisc.ac.uk/publications/briefingpapers/2006/pub\\_digipreservationbp.aspx](http://www.jisc.ac.uk/publications/briefingpapers/2006/pub_digipreservationbp.aspx) accessed 08.10.2020.
- Pillen, D., Eckard, M. (2022). The impact of the shift to cloud computing on digital recordkeeping practices at the University of Michigan Bentley historical library. *Arch Sci* . <https://doi.org/10.1007/s10502-022-09395-2>
- Pinfield, S. & James, H. (2003), The Digital Preservation of e-Prints. *D-Lib Magazine*, 9(9) available at: <http://www.dlib.org/dlib/september03/pinfield/09pinfield.html>
- Plath, Debbie. (2013). Evidence-based practice. In *Social Work Theories and Methods*, 2nd ed. Edited by Mel Gray and Stephen Webb. London: Sage, pp. 229–40.
- Powell, R. (2006). Evaluation Research: An Overview. *Library Trends*, 55, 102-120. Retrieved from: <http://hdl.handle.net/2142/3666> on 03/06/2020
- Property Records Industry Association (PRIA). (2017). TIFF-PDF/A – A White Paper. Retrieved from: <https://docplayer.net/194067446-Tiff-pdf-a-white-paper.html> on 30/11/2020
- Pryor, G. & Donnelly, M. (2009). Skilling up to do data: whose role, whose responsibility, whose career? *The International Journal of Digital Curation*, 4(2), 158-170. <https://doi.org/10.2218/ijdc.v4i2.105>
- Qu, S.Q. & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting & Management*, 8(3), 238-264. <https://doi.org/10.1108/11766091111162070>

- Quisbert, H. (2006). A Framework for the Development of Archival Information Systems, [ Unpublished masters thesis] Luleå University of Technology. Retrieved from: <https://www.diva-portal.org/smash/get/diva2:990983/FULLTEXT01.pdf> on 10/06/2021
- Rahmanto K.N. & Riasetiawan, M. (2018). *Data Preservation Process in Big Data Environment using Open Archival Information System*, 4th International Conference on Science and Technology (ICST), 2018, pp. 1-5, doi: 10.1109/ICSTC.2018.8528669.
- Rajasekar, S., Pitchai, P. N. & Veerapadran, C. (2006). Research Methodology. Retrieved from: <https://arxiv.org/abs/physics/0601009> on 02/06/2019
- Rauch, C. et al, (2007). File-formats for preservation: evaluating the long-term stability of file-formats. *ELPUB*,101-106 retrieved from: [https://elpub.architexturez.net/system/files/pdf/122\\_elpub2007.content.pdf](https://elpub.architexturez.net/system/files/pdf/122_elpub2007.content.pdf) on 06/06/2021
- Ravenwood, C. Muir, A & Matthews, G. (2015). Stakeholders in the Selection of Digital Material for Preservation: Relationships, Responsibilities, and Influence. *Collection Management*, 40:2, 83-110, DOI: 10.1080/01462679.2015.1011816
- Registry of Open Access Repository Mandates and Policies (ROARMAP). Accessed at: <http://roarmap.eprints.org/> on 01/02/2022
- Rescher, N. (2000). *Realistic Pragmatism: An Introduction to Pragmatic Philosophy*. SUNY Press, Albany.
- Research Libraries Group. (2002). Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report. Mountain View, CA. Retrieved from: <http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf?urlm=161690> on 01/06/2020
- Richards, K. (2003). *Qualitative Inquiry in TESOL*. Basingstoke: Palgrave Macm
- Rieger, O. Y., Schonfeld, R. C., & Sweeney, L. (2022). The Effectiveness and Durability of Digital Preservation and Curation Systems. A research report. *ITHAKA*. <https://doi.org/10.18665/sr.316990>
- Rimkus, K., Padilla, T., Popp, T. & Martin, G. (2014). Digital preservation file format policies of ARL member libraries: an analysis. *D-Lib Magazine*, 20(3/4). Retrieved from <http://www.dprylib.org/dlib/march14/rimkus/03rimkus.html> on 14/05/2019
- Rinehart, A. K., Prud'homme, P. A., & Huot, A. R. (2014). Overwhelmed to action: digital preservation challenges at the under-resourced institution. *OCLC Systems and Services*, 30(1), 28-42. doi:10.1108/OCLC-06-2013-0019
- Robertson, W. C. & Borchert C. A. (2014), Preserving Content from Your Institutional Repository. *Serials Librarian*, 66, 278–288. DOI: 10.1080/0361526X.2014.881209

- Robins, D. (2002), From virtual libraries to digital libraries: the role of digital libraries in information communities' in C. F. Thomas, Marcel Dekker (eds) *Libraries, the Internet and Scholarship: Tolls and Trends Converging*, New York, pp. 45-76
- Ross, S. & Hedstrom, M. (2005). Preservation research and sustainable digital libraries. *International Journal on Digital Libraries*, 5(4), 317-324. DOI: <http://dx.doi.org/10.1007/s00799-004-0099-3>
- Roy, B. K., Biswas, S. C., & Mukhopadhyay, P. (2012). An analytical study of institutional digital repositories in India. *Library Philosophy and Practice*, 1(1), 1-13. Retrieved from: <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1766&context=libphilprac> on 06/06/2021
- Roy, B., Biswas, S., & Mukhopadhyay, P. (2018). Towards Open Access Self Archiving Policies: A Case Study of COAR. *LIBER Quarterly*, 28(1), 1–33. <https://doi.org/10.18352/lq.10227>
- Rumsey, S. (2006). The purpose of institutional repositories in UK higher education: a repository manager's view [online]. London: LSE Research Online. Retrieved from: <http://eprints.lse.ac.uk/archive/00000800> on 02/06/2020
- Ruusalepp, R., & Dobрева, M. (2012). Digital preservation services: state of the art analysis. Retrieved from: <https://www.um.edu.mt/library/oar/handle/123456789/311> on 06/06/2021
- Ryan, G. (2018). Introduction to positivism, interpretivism and critical theory. *Nurse Researcher* 25(4). 41–49. <https://doi.org/10.7748/nr.2018.e1466>
- Ryan, H. (2014). Who's afraid of file format obsolescence? Evaluating file format endangerment levels and factors for the creation of a file format endangerment index. [Unpublished doctoral dissertation]. University of North Carolina at Chapel Hill. Retrieved from: <https://doi.org/10.17615/fhkk-dv90>
- Sandy, H. M. & Corrado, E. M. (2017). Bringing Content into the Picture: Proposing a Tri-Partite Model for Digital Preservation. *Journal of Library Administration*, 58(1), 1-17. <https://doi.org/10.1080/01930826.2017.1385988>
- Saunders, M., Lewis, P. & Thornhill, A. (2007). *Research Methods for Business Students*. (4<sup>th</sup> Ed.). London: Prentice Hall.
- Sekaran, U. (2003). *Research Methods for Business: A Skill-Building Approach*. (4<sup>th</sup> Ed). John Wiley & Sons, New York.
- Shajitha, C. & Abdul K.C. (2021), Institutional repositories in South India: an exploratory study. *Global Knowledge, Memory and Communication*. DOI 10.1108/GKMC-07-2020-0099
- Shala, L. & Shala, A. (2016). File Formats - Characterization and Validation. *IFAC-Papers OnLine*, 49(29), 253-258. <https://doi.org/10.1016/j.ifacol.2016.11.062>



- Shearer, K. (2006), The CARL institutional repositories project: A collaborative approach to addressing the challenges of IRs in Canada. *Library Hi Tech*, 24 (2), 165-172. <https://doi.org/10.1108/07378830610669547>
- Shilling, S. (2018). Collecting and preserving digital materials. Retrieved from <https://www.history.org.au/wp-content/uploads/2018/12/Collecting-and-Preserving-Digital-Materials.pdf> on 09/12/2020
- Shiloba G. E. & Mohammed, Z. (2013), Digital preservation policy in National Information Centres in Nigeria. *The Electronic Library*, 31(4), 483-492. DOI <https://doi.org/10.1108/EL-01-2012-0011>
- Simon, A. & Kiszl, P. (2021). Aspects of the Long-Term Preservation of Digitized Catalogue Data: Analysis of the Databases of Integrated Collection Management Systems Preservation, *Digital Technology & Culture*, 50(2) pp. 51-64. <https://doi.org/10.1515/pdte-2021-0007>
- Sinclair, P. et al. (2011). Are you Ready? Assessing Whether Organisations are Prepared for Digital Preservation. *The International Journal of Digital Curation*, (6)1: 268-281. DOI: 10.2218/ijdc.v6i1.187
- Singh, Y.K. (2006). *Fundamental of research methodology and statistics*. New Delhi: New Age International
- Skinner, K. & Halbert. M. (2009). The MetaArchive Cooperative: A Collaborative Approach to Distributed Digital Preservation. *Library Trends*. 57 (3): 371–392. Retrieved from: <http://hdl.handle.net/2142/1360> on 10/12/2020
- Smith, K. (2008). Institutional repositories and e-journal archiving: What are we learning? *The Journal of Electronic Publishing*, 11(1), 1-10. DOI: <https://doi.org/10.3998/3336451.0011.107> on 08/05/2020
- Solomon D., Soyemi, O.D (2021). ICT competence and digital preservation practices of library personnel in selected Nigerian private universities. *Inter. J. Acad. Res. Educ. Rev.* 9(5): 248-259 doi: 10.14662/Ijarer2021.195
- Srirahayu et al (2020). Readiness for Digital Preservation in Indonesia. *Library Philosophy and Practice (e-journal)*. 4625. <https://digitalcommons.unl.edu/libphilprac/4625>
- Stern, E (2004) Philosophies and types of evaluation research. In *The Foundations of Evaluation and Impact Research. Third report on vocational training research in Europe: background report*. Luxembourg: Office for Official Publications of the European Communities (Cedefop Reference Serie, 58), 12–42. Retrieved from: [https://www.cedefop.europa.eu/files/BgR1\\_Stern.pdf](https://www.cedefop.europa.eu/files/BgR1_Stern.pdf) on 17/02/2021
- Strodl, S., Becker, C., Neumayer, R., & Rauber, A. (2007). *How to choose a digital preservation strategy: Evaluating a preservation planning procedure*. In Proceedings of the 2007 Conference on Digital Libraries. Vancouver, BC, Canada, June 18 - 23, 2007. JCDL '07. ACM: New York, 29-38. Retrieved from:

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.5970&rep=rep1&type=pdf> on 04/05/2019

- Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *International Journal of Advance Research in Management*, 5(2), 18-27. Retrieved from: <https://hal.archives-ouvertes.fr/hal-02546796/document> on 24/02/2021
- Talboom, L & Underdown, D. (2019). Access is What we are Preserving: But for Whom? Digital Preservation Coalition (DPC) Blog. Access from <https://www.dpconline.org/blog/access-what-we-are-preserving> on 19/10/2021.
- Tallman, N. & Work, L (2018, September 24-28). *Approaching Appraisal Guidelines and Criteria to Select for Digital Preservation*. International Conference on Digital preservation, Boston, MA, United States retrieved from: <https://phaidra.univie.ac.at/detail/o:923633.pdf> on 03/06/2019
- Tapscott, L. (2019). I'm Gonna Live Forever! Assessing the Preservation of Digital Collections, <https://alcts.ala.org/news/2019/assessing-the-preservation-of-digital-collections/> on 06/06/2021
- Termens, M., Ribera, M. and Locher, A. (2015), An analysis of file format control in institutional repositories. *Library Hi Tech*, 33 (2), pp. 162-174. doi: 10.1108/LHT-10-2014-0098
- The British Library (2013). *Building a Preservation Policy*. London: The Preservation Advisory Centre. Retrieved from: [file:///C:/Users/ADMIN/Downloads/building-a-preservation-policy%20\(3\).pdf](file:///C:/Users/ADMIN/Downloads/building-a-preservation-policy%20(3).pdf) on 03/06/2019
- The Commission on Preservation and Access and the Research Libraries Group (1996). Preserving Digital Information. *Report of the Task Force on Archiving of Digital Information*. Retrieved from: <https://www.oclc.org/content/dam/research/activities/digpresstudy/final-report.pdf> on 08/01/2021
- The UK National Archives (2011). Digital preservation policies guidance for archives. Retrieved from: <http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf> on 02/06/2020
- Thibodeau K. (2002). *Overview of technological approaches to digital preservation and challenges in coming years*. In: The State of Digital Preservation: An International Perspective – Conference Proceedings, Documentation Abstracts, Inc., Institutes for Information Science, Washington, D.C., April 24-25, 2002. Council on Library and Information Resources, Washington, D.C., <https://www.clir.org/pubs/reports/pub107/thibodeau/>
- Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository. *Journal of Digital Information*, 8(2), 1-5. Retrieved from: <https://journals.tdl.org/jodi/index.php/jodi/article/view/197/174> on 03/05/2020

- Tieman J. (2015). Trustworthy Digital Repositories. Us government publishing office. Retrieved from: <https://www.fdlp.gov/file-repository/fdlp-academy-video-supplemental-materials/3236-trustworthy-digital-repositories-slides/file> on 16.02.2021
- Tomasok J. (2018). Design and implementation of Archival Storage component of OAIS Reference Model. [Unpublished Masters thesis] Masaryk University, retrieved from: <https://is.muni.cz/th/ykhgv/tomasek-diploma-thesis.pdf> on 04/06/2021
- Tombs, M. & Pugsley, L. (2020), How to Understand Research Philosophies and Paradigms in Medical Education. Cardiff University. available at: Retrieved from: [https://www.cardiff.ac.uk/\\_data/assets/pdf\\_file/0003/2425170/How-to-Understand-Research-Philosophies.pdf](https://www.cardiff.ac.uk/_data/assets/pdf_file/0003/2425170/How-to-Understand-Research-Philosophies.pdf) on 22/02/2021
- Traczyk, T, et al (eds) (2017). *Digital Preservation: Putting It to Work*. Warsaw, Poland; Springer International Publishing
- Trehub, A. et al. (2019). LOCKSS Networks: Community-Based Digital Preservation. In *Digital Preservation in Libraries: Preparing for a Sustainable Future*, edited by Jeremy Myntti and Jessalyn Zoom, 217–240. Association for Library Collections & Technical Services Monograph. Chicago, IL: American Library Association. Retrieved from: <https://aurora.auburn.edu/handle/11200/49351> on 06/06/2021
- Trehub, A., & Halbert, M. (2012). *Safety in Numbers: Distributed Digital Preservation Networks*. In Proceedings of the IFLA World Library and Information Congress, 78th IFLA General Conference and Assembly, Helsinki, Finland. Retrieved from: <https://www.ifla.org/past-wlic/2012/216-trehub-en.pdf> on 10/12/2020
- Umana, S.P. (2020). Digital preservation in institutional repositories in Namibian academic libraries: a case study of the Namibia University of Science and Technology (NUST) and the university of Namibia (UNAM). [https://repository.unam.edu.na/bitstream/handle/11070/2908/umana\\_2020.pdf?squence=1&isAllowed=y](https://repository.unam.edu.na/bitstream/handle/11070/2908/umana_2020.pdf?squence=1&isAllowed=y)
- United Nations Educational, Scientific and Cultural Organization (2013). Guidelines for the preservation of national heritage. Retrieved from: <http://wayback.archive-it.org/10611/20160802151228/http://unesdoc.unesco.org/images/0013/001300/130071e.pdf> on 02/06/2019
- University of Nairobi (2022). University of Nairobi institutional repository retrieved at: <http://erepository.uonbi.ac.ke/> on 10/12/2022
- University of Nairobi. (2012). Open access policy, retrieved from <https://uonlibrary.uonbi.ac.ke/node/1482> on 12/08/2017
- Van, Garderen, P. et al, (2012). *The Archivematica project: meeting digital continuity's technical challenges*. The memory of the world in the digital age: digitization and preservation. held on 26–28. September 2012 in Vancouver, British Columbia, Canada

- Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterizing and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, 18(1), 1 - 18. DOI: 10.1186/s12874-018-0594-7
- Verdegem, R. & Slats, J. (2004), Practical experiences of the Dutch digital preservation test-bed, *VINE*, 34(2), 56-65. <https://doi.org/10.1108/03055720410531004>
- Vij, R. (2017). Digital Preservation of Information Resources: Issues and Strategies. *International Journal of Research in Humanities & Soc. Sciences*. 5(5) [https://www.raijmr.com/ijrhs/wp-content/uploads/2017/11/IJRHS\\_2017\\_vol05\\_issue\\_05\\_01.pdf](https://www.raijmr.com/ijrhs/wp-content/uploads/2017/11/IJRHS_2017_vol05_issue_05_01.pdf)
- Visser, M. (2019). Pragmatism, critical theory and business ethics: Converging lines. *Journal of Business Ethics*, 156, 45–57. doi: 10.1007/s10551-017-3564-9
- Vrana, R. (2011). Digital repositories and the future of preservation and use of scientific knowledge. *Informatol*. 44(1). 55-62. Retrieved from: [https://bib.irb.hr/datoteka/510540.DIGITALNI\\_REPOZITORIJI\\_I\\_BUDUCNO\\_ST\\_OcUVANJA.pdf](https://bib.irb.hr/datoteka/510540.DIGITALNI_REPOZITORIJI_I_BUDUCNO_ST_OcUVANJA.pdf) on 03/06/2019
- Waddington, S., Zhang, J., Knight, G., Jensen, J., Downing, R., & Ketley, C. (2013). Cloud repositories for research data – addressing the needs of researchers. *Journal of Cloud Computing: Advances, Systems and Applications*, 2(1), 13. <https://doi.org/10.1186/2192-113X-2-13>.
- Wakeling, S. et al, (2019). Academic communities: The role of journals and open-access mega-journals in scholarly communication, *Journal of Documentation*, 75(1) pp. 120-139. <https://doi.org/10.1108/JD-05-2018-0067>
- Waller, M. & Sharpe, R. (2006). *Mind the Gap: Assessing Digital Preservation Needs in the UK*, Digital Preservation Coalition, York. Retrieved from: [www.dpconline.org/docs/reports/uknamindthegap.pdf](http://www.dpconline.org/docs/reports/uknamindthegap.pdf). on 06/06/2019
- Walliman, N. (2011). *Research Methodology*. London: Routledge
- Wang, H. (2017). Where the Rubber Meets the Road: Self-Audit in University Digital Repositories. [Unpublished Masters thesis] University of North Carolina. Retrieved from: <https://core.ac.uk/download/pdf/210609334.pdf> on 18/02/2021
- Waters, D. & Garrett, J. (eds.) (1996). *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*. Washington, DC: Commission on Preservation and Access and the Research Libraries Group. Retrieved from: <https://www.clir.org/pubs/reports/pub63/> on 06/04/2018
- Weatherburn, J. (2017). Collaboration and Digital Preservation. <https://jayeweatherburn.com/2017/12/18/collaboration-and-digital-preservation/>

- Webb, C., Pearson, D. & Koerbin, P (2013) 'Oh, you wanted us to preserve that?!' Statements of Preservation Intent for the National Library of Australia's Digital Collections. *D-Lib Magazine* 19(1/2). Retrieved from: <http://www.dlib.org/dlib/january13/webb/01webb.html> on 04/08/2020
- Weinraub, E. et al. (2018). Beyond the Repository: Integrating Local Preservation Systems with National Distribution Systems. Report. <https://doi.org/10.21985/N28M2Z>
- Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21-33. <https://doi.org/10.1177/109821409801900103>
- White M & Marsh E (2006) Content analysis: a flexible methodology. *Library Trends* 55(1):22–45. DOI: 10.1353/lib.2006.0053
- Whitt, R.S. (2017). Trough A Glass, darkly: Technical, Policy, and Financial Actions to Avert the Coming Digital Dark Ages. *Santa Clara High Technology Law Journal*, 33(2). Retrieved from: <https://assets.pubpub.org/auraylo2/11521405338888.pdf> on 01/12/2019
- Wilczek, E. & Glick K. (2006) *Fedora and the Preservation of University Records Project: Reports and Findings*, Tufts University and Yale University, Final Narrative Report to National Historical Publications and Records Commission, September 27.
- Wiles, R. (2013) *What Are Qualitative Research Ethics?* London: Sage.
- Williamson, K. & Johanson, G. (Eds.) (2017). *Research methods: information, systems, and context* (2<sup>nd</sup> ed.). Cambridge, MA: Chandos
- Wilson, V. (2014). Research Methods: Triangulation. *Evidence Based Library and Information Practice*, 9(1), 74–75. <https://doi.org/10.18438/B8WW3X>
- Wilson, V. (2016). Research Methods: Content Analysis. *Evidence-Based Library and Information Practice*, 11(41-43). <https://doi.org/10.18438/B8CG9D>
- Wittenberg, C., et al (2018). Challenges and opportunities in the evolving digital preservation landscape: reflections from Portico. *Insights*. 31, 28. DOI: <http://doi.org/10.1629/uksg.421>
- Woodyard, D. (2004), Significant property: digital preservation at the British library VINE: *The Journal of Information and Knowledge Management Systems*, 34(1), 17-20. DOI: 10.1108/03055720410530942
- Yakel, E. et al (2008). Institutional repositories and the institutional repository. *American Archivist*, 71(2), 323-349. Retrieved from: <http://hdl.handle.net/2027.42/106421> on 10/10/2019
- Yannis, T. & Yannis, M. (2018). *Cinderella's Stick: A Fairy Tale for Digital Preservation* (1st. ed.). Springer Publishing Company, Incorporated.

- Yin R. K. & Campbell D. T. (2018). *Case study research and applications: design and methods* (Sixth). SAGE Publications.
- Yin, R. K. (2009). *Case study research: Design and methods* (5<sup>th</sup> ed.). Thousand Oaks; Sage.
- Zaidman- Zait A. (2014) Content Analysis. In A. Michalos (Ed.) *Encyclopedia of Quality of Life and Well-Being Research*. Springer Netherlands, Dordrecht, pp. 1258-1261
- ZBW-Leibniz Information Centre for Economics (n.d). Preferred file formats in the Digital Archive of the ZBW | Retrieved 21/2/2020 , from <https://www.zbw.eu/en/about-us/key-activities/digital-preservation/file-formats>
- Zierau, Eld. (2017). OAIS and Distributed Digital Preservation in Practice. An exploration of Danish and other use cases that contributed to the development of the Outer OAIS–Inner OAIS Model for Distributed Digital Preservation. 14th International Conference on Digital Preservation iPRES, Kyoto, Japa. <https://doi.org/10.5281/zenodo.1308845>
- Zikmund, W. et al (2012). *Business Research Methods*. (9<sup>th</sup> ed.) Ohio; South-Western, Cengage Learning
- Zuccala, A. et al (2006). Digital Repository Management Practices, User Needs and Potential Users: An Integrated Analysis. Joint Information Systems Committee Project. Report retrieved from: <http://cybermetrics.wlv.ac.uk/DigitalRepositories/FinalReport.pdf> on 14/08/2019
- Zuccala, A., Oppenheim, C. & Dhiensa, R. (2008). Managing and evaluating digital repositories. *Information Research*, 13(1) paper 333. Retrieved from: <http://InformationR.net/ir/13-1/paper333.html> on 10/10/2019
- Žukauskas, P. et al (2018). Philosophy and Paradigm of scientific research. *Management Culture and Corporate Social Responsibility*, 121. DOI: 10.5772/intechopen.70628

## APPENDIX 1: INTERVIEW SCHEDULE

### A. Characteristics of materials received in the repository

1. What types of digital resources are accepted in the digital repository in regard to content?
2. Of the above which type of resource has the largest content?
3. Are the boundaries for the digital collection clearly defined in an acquisition policy?
4. What formats of information resources are accepted in the repository?
5. Are these formats clearly articulated in a policy?
6. Are their guidelines for depositors and ingest?
7. What type of metadata is collected during ingest?
8. Is the metadata collected enough for long term preservation?
9. What metadata standards do you apply?

### B. Preservation Policy

Policies in place to promote preservation for long-term access

1. Does your repository have a digital preservation policy?
2. If yes, does the repository adhere to it?
3. Has the policy been revised and by whom?
4. Who is responsible for the policy?
5. Does the policy address rights management issues in relation to digital preservation?
6. If no to 1 above, do you plan to have a digital preservation policy?
7. If yes, what is your time line?

### C. Digital preservation planning

1. Does the repository have a digital preservation plan?
2. If yes does it spell out the responsibilities for preservation?
3. Is the preservation done in-house or the services are outsourced
4. Does the plan provide for financial sustainability?

5. Does the plan provide for technology sustainability and technology monitoring?
6. Does the plan provide for skills development?
7. Are there mechanisms in place to change the preservation plan, as user needs and technology needs change?

D. Digital preservation strategies

1. What preservation strategies are in place?
2. Are they well-articulated in the preservation plan and policy?
3. Does the repository have in place procedures and templates for ensuring information is captured when accessioning new digital records such as: file formats, bitstream checks, software used for creating and editing including version, operating system used, compression formats for multimedia, others?
4. What challenges are you experiencing in preservation?
5. What solutions would you offer for the above challenges?



**APPENDIX 2: OBSERVATION AND DOCUMENT ANALYSIS GUIDE**

1. What resources are accepted in the repository? (Text, audio, video)
2. What formats are accepted (pdf, jpeg, mpeg among others)?
3. Is the intention of long-term preservation well-articulated in the policy?
4. Is the metadata collected sufficient to support long-term access to the resources?
5. Metadata schemes used for long-term preservation
6. Suitability of repository platform to support longterm preservation (File format management, integrity checks; preservation metadata)
7. Do the repository contracts articulate well on intellectual property right issues?
8. Documentation on actions already taken for long-term access
9. Documentation on responsibilities for technology watch
10. Is the policy clear of the review cycle?
11. Are decisions on preservation actions well documented?
12. Is budgeting done with digital preservation in mind?

## APPENDIX 3: NACOSTI RESEARCH AUTHORIZATION



### NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY AND INNOVATION

Telephone: +254-20-2213471,  
2241349, 3310571, 2219420  
Fax: +254-20-318245, 318249  
Email: dg@nacosti.go.ke  
Website: www.nacosti.go.ke  
When replying please quote

9<sup>th</sup> Floor, Utalii House  
Uhuru Highway  
P.O. Box 30623-00100  
NAIROBI-KENYA

Ref. No. **NACOSTI/P/17/88668/17178**

Date: **27<sup>th</sup> June, 2017**

Hellen Mukiri Ndegwa  
Moi University  
P.O. Box 3900-30100  
**ELDORET.**

#### **RE: RESEARCH AUTHORIZATION**

Following your application for authority to carry out research on *“An evaluation of digital preservation practices in institutional repositories (IRS) of selected public universities in Kenya,”* I am pleased to inform you that you have been authorized to undertake research in **Kiambu and Nairobi Counties** for the period ending **22<sup>nd</sup> June, 2018.**

You are advised to report to **the Vice Chancellors of the selected public Universities, the County Commissioners and the County Directors of Education, Kiambu and Nairobi Counties** before embarking on the research project.

On completion of the research, you are expected to submit **two hard copies and one soft copy in pdf** of the research report/thesis to our office.

  
GODFREY P. KALERWA MSc., MBA, MKIM  
FOR: DIRECTOR-GENERAL/CEO

Copy to:

The Vice Chancellors  
Selected Universities


The County Commissioner  
Kiambu County.

### APPENDIX 4: NACOSTI RESEARCH PERMIT


**THIS IS TO CERTIFY THAT:**  
**MS. HELLEN MUKIRI NDEGWA**  
**of MOI UNIVERSITY, 1957-10101**  
**KARATINA, has been permitted to**  
**conduct research in *Kiambu , Nairobi***  
**Counties**


**on the topic: AN EVALUATION OF**  
**DIGITAL PRESERVATION PRACTICES IN**  
**INSTITUTIONAL REPOSITORIES (IRS) OF**  
**SELECTED PUBLIC UNIVERSITIES IN**  
**KENYA**

**for the period ending:**  
**22nd June,2018**

  
.....  
**Applicant's**  
**Signature**

**Permit No : NACOSTI/P/17/88668/17178**  
**Date Of Issue : 27th June,2017**  
**Fee Recieved :Ksh 2000**



  
.....  
**Director General**  
**National Commission for Science,**  
**Technology & Innovation**

## APPENDIX 5: UON AUTHORITY TO COLLECT DATA



**UNIVERSITY OF NAIROBI**  
**OFFICE OF THE DEPUTY VICE - CHANCELLOR**  
 (Research, Production & Extension)  
 Prof. Lucy W. Irungu B.Sc., M.Sc., Ph.D.

P.O. Box 30197-GPO,  
 00100, Nairobi-Kenya  
 Telephone: +254-20-2315416 (DI), 318262

Fax: 0202317251  
 Email: dvrpe@uonbi.ac.ke

UON/RPE/3/5

July 4, 2017

Hellen M. Ndegwa,  
 Moi University,  
 P.O. Box 3900-30100,  
 ELDORET.

Dear Ndegwa,

**AUTHORITY TO COLLECT DATA AT THE UNIVERSITY OF NAIROBI**

I refer to your request dated June 28, 2017 to collect data at the University of Nairobi for your PhD research entitled: *"Evaluation of Digital Preservation Practices in Institutional Repositories (IRs) of selected Public Universities in Kenya"*.

I write to inform you that your request has been approved.

You are however required to share the findings of your study with the University of Nairobi by depositing a copy of your research findings with the Director, Library and Information Services on completion of your study.

**LUCY W. IRUNGU**  
**DEPUTY VICE-CHANCELLOR**  
**(RESEARCH, PRODUCTION AND EXTENSION)**  
 &  
**PROFESSOR OF ENTOMOLOGY**

Copy to: Vice-Chancellor  
 DVC, A&F  
 DVC, AA  
 DVC, SA  
 Director, Library & Information Services  
 Registrar Administration



ISO 9001:2008 CERTIFIED

*The Fountain of Knowledge Providing leadership in academic excellence*

## APPENDIX 6: JKUAT APPROVAL TO CARRY OUT RESEARCH



**JOMO KENYATTA UNIVERSITY  
OF  
AGRICULTURE AND TECHNOLOGY**

**RESEARCH PRODUCTION & EXTENSION DIVISION  
(OFFICE OF THE DIRECTOR RESEARCH)**

P. O. Box 62000-00200 Nairobi, Kenya Tel 067587022. Email: [research@rpe.jkuat.ac.ke](mailto:research@rpe.jkuat.ac.ke)

Hellen M. Ndegwa,  
Moi University,  
P. O. Box 3900-30100,  
Eldoret

**REF: APPROVAL TO CARRY OUT RESEARCH – HELLEN NDEGWA**

Following your request to collect data from the University Library for your doctoral study titled: 'An evaluation of digital preservation practices in institutional repositories (IRs) of selected public universities in Kenya', I am pleased to inform you that your request has been approved.

You are advised to report to the University Librarian before starting data collection.

On completion of the work, you are expected to deposit a copy of the thesis to the University Librarian.

A handwritten signature in blue ink, appearing to read 'P. Mbindyo'.

**Dr. Patrick Mbindyo**  
**DIRECTOR RESEARCH**

Copy to: University Librarian



*JKUAT is ISO 9001:2008 and 14001:2004 Certified*  
*Setting Trends in Higher Education, Research and Innovation*

## Appendix 7: Similarity Index

SR195



*ISO 9001:2019 Certified Institution*

**EDU 999 THESIS WRITING COURSE**

*PLAGIARISM AWARENESS CERTIFICATE*

This certificate is awarded to

**HELLEN MUKIRI NDEGWA**

**IS/DPHIL/01/12**

In recognition for passing the University's plagiarism  
Awareness test for the thesis: **EFFICACY OF DIGITALPRESERVATION PRACTICES IN  
INSTITUTIONAL REPOSITORIES (IRS) OF SELECTED PUBLIC UNIVERSITIES IN KENYA**  
with a similarity index of 7% and striving to maintain academic integrity.

Awarded by:



**Prof. Anne Syomwene Kisilu**  
CERM-ESA Project Leader    Date: 15/06/2023