Digital preservation and institutional repositories: case study of universities in Kenya

Irene Moseti (PhD)
1st affiliation: Institution: Moi University, School of Information Sciences, Department of Information Technology.

Second affiliation: Research Collaborator: InterPARES Africa Team, coordinated by the Department of Information Science, University of South Africa
irenemorara@gmail.com

Abstract

This article reports on part of the findings of a study that was completed at the University of KwaZulu Natal in 2015. The study investigated the strategies used by universities in Kenya for the preservation of their scholarly content and is the first of its kind. The survey method was used within a multiple case study design. The data was collected using questionnaires administered to 350 academic staff and 370 postgraduate students conveniently selected from six universities in Kenya. Personal interviews were used to collect data from the university librarians in these universities. It was found that scholars at the universities were personally engaged in preservation of their digital information but did not extensively use university digital archives, servers or repositories. This was largely attributed to lack of awareness of the important role of digital repositories in digital preservation. The present study reveals that even with the existence of institutional repositories, much need to be done to create more awareness and acceptance of digital repositories as well as integrating digital preservation best practices into the daily repository management activities.

Key words: digital preservation, institutional repositories, scholarly content, universities, Kenya

1. Introduction

With the advances of technology, an increasing number of Higher Education Institutions are implementing digital repositories aimed at helping these institutions to develop coherent and coordinated approaches to capture, identify, store and retrieve intellectual assets such as datasets, course material and research papers (Hoorens et al., 2008). Also known as Institutional Repositories (IR), these are digital collections which capture and preserve the intellectual output of a single or multi-university community. Usually IRs adopt OA principles which facilitate free and unlimited access to the content (Crow, 2002). IRs have become one of the fastest growing elements of the digital library genre due to their potential to reform the current system of scholarly communication and their role in advancing the open access movement. As a result, many academic libraries, especially in research universities, have invested human and technical resources to build a robust technical infrastructure that will foster access to the intellectual, cultural, and administrative output of their institutions. The hope is to gain enhanced access to faculty research and increased visibility of research generated within the university (Jantz and Wilson, 2008, p. 187).

Several authors note that a key objective of IR is to provide long term preservation of digital materials in addition to organization and access or distribution (Hockx-Yu, 2006, Lynch, 2003). However, Hockx-Yu (2006) observes that the focus of many repository activities has been on creating repositories, promoting discovery and access and/or encouraging the necessary cultural change. Digital preservation has not been embedded as an integral part of the repositories' workflow and there is neither much experience nor commonly agreed best practice as to how digital preservation is best performed.

Digital preservation is the set of processes and activities that ensures long-term, sustained storage of, access to and

interpretation of digital information. Digital preservation activities aim to keep valuable and useful digital material available for future generations of scholars, researchers and other users. Curation is an applied form of preservation that focuses on interpretation and is often used in relation to working with scientific datasets (JISC, 2014). According to Brandt (2007, p. 366), "in the Information Age, curation can be defined as essential activities and systems that facilitate access, dissemination, and archiving of e-research. It includes protocols and tools that provide descriptive analyses of digital collections and objects to augment discovery, management, use, reuse, and preservation. Curation is about policies and consultation, as well as tools and systems".

Lord et al., (2004) observe that modern instruments and computing resources have enabled e-research and a new order of collaborating and inter-disciplinary research. This has increased access to collections of primary research data and information. However, the same technology tools put the data created at risk, raising serious and complex issues of strategy, policy and practice regarding the creation, management, and long-term curation of the data. Lord et al. argue that much needs to be done to enable this data to remain available and valid to future researchers. Lord et al. also note that e-Science curation entails three key activities: *curation* (managing and promoting the use of data from its point of creation, ensuring it is fit for contemporary purpose, and available for discovery and reuse); *archiving* (ensures that data is properly selected, stored, accessible and that its logical and physical integrity is maintained over time); and *preservation* (specific items of data are maintained over time so that they can still be accessed and understood through changes in technology).

2. Literature review

According to Brown et al. (2009), the term 'digital preservation' is used to refer to the overall approach to preserving information and records created using computers, including electronic records. The authors describe the fundamental issues in digital preservation including: the characteristics of electronic records and the fact that they are composed of different digital objects; the role of different software programs for identifying the characteristics of those different digital objects that make up electronic records; the difference between active and passive preservation of electronic records; and different types of preservation, including refreshing data, replicating data, migration and emulation.

Banach and Li (2011) argue that in the digital age, preserving information has become a more complex task because digital information is fragile and faces many threats including technological obsolescence and the deterioration of digital storage media. Conway (1996) points to a central dilemma of the digital age: "Our capacity to record information has increased exponentially over time while the longevity of the media used to store the information has decreased equivalently." However, Lavoie and Dampsey (2004) argue that over time, the focus of digital preservation has shifted away from the need to take immediate action to "rescue" threatened materials, and instead moved toward the realization that perpetuating digital materials over the long-term involves the observance of careful digital asset management practices diffused throughout the information lifecycle. Further, Lavoie and Dampsey (2004) contend that digital preservation techniques are most effective when they are pre-emptive since it is often impossible or prohibitively expensive to restore a digital object that has become corrupted or obsolete. The authors therefore advise that digital preservation within institutions responsible for collecting digital content ought to be integrated as a continuous process within the day to day management of this content.

Pennock (2007) advocates for digital curation which is the active management

and appraisal of digital information over its entire life cycle. She opines that approaching digital information management from a life cycle perspective facilitates continuity of service and supports verification of the origin of digital data despite technological and organisational changes in their context. To succeed however, digital curation requires significant input and buy-in from the range of stakeholders with responsibilities for the materials at different stages of the life cycle. Stakeholders range from creators of digital material to curators such as librarians and archivists, IT staff, and other stakeholders, including management.

The growing awareness of the urgency of digital preservation has led to the development of various approaches that deal with the question of preserving digital objects over long periods of time. Current approaches can be divided into *migration* which transforms the objects to more widely accessible representations and *emulation* which creates a technical environment where the objects can be rendered or performed (Becker et al., 2009). However, Becker et al (2009) caution that both approaches face challenges for more exotic and complex compound objects, as well as large amounts of data. Recently, there has been emphasis on the use of open formats to facilitate access to digital data in the long term. According to PC.net (2014), a file format describes the way data is stored in a file. It defines the data structure (how the data is organised in the file) as well the type of data that the file contains. Some file formats are "open formats," meaning they are publicly available and all software developers can use them. Other file formats are proprietary or "closed formats," meaning that only specific applications can open them. There are specific formats for images (for example, JPEG, PNG, GIF, TIF, BMP), simple text (ASCII, often marked with the .txt extension), for formatted text (HTML, RTF, DOC) and for printer-ready documents (PDF, PS).

3. Research objective

Much has been written on the data deluge and information overload facing modern scientists. Bell et al.(2009) attribute this to developments in science that have enabled generation and analysis of hundred-to-thousand-fold increases in data volumes from satellites, telescopes, and supercomputers. Added to this are data management challenges emanating from "born digital" data in files, spreadsheets or databases stored on hard drives, digital notebooks, web sites, blogs and wikis. Management, curation, and archiving of these digital data and information are becoming increasingly burdensome for research scientists. As Jackson (2012) states, digital resources will not survive or remain accessible by accident: pro-active preservation is needed. Lord et al. (2004) found that awareness of long-term data curation was generally low among researchers, thus they needed encouragement to engage more in the curation of their own data. Marshall et al, (2006) noted that experienced home computer users are creating, receiving and finding an increasing number of digital belongings but they have already lost irreplaceable digital artefacts such as photos, creative efforts and records. Participants in their study used strategies such as backup and file replication for digital safekeeping but were unable to implement them consistently. From a study in South Africa, Groenewald and Breytenbach (2011) indicate a similar lack of knowledge on preservation strategies and the management of digital objects on personal computers, as well as a need for training in basic digital preservation methods.

Western University in Canada (2016) state that digital preservation ensures the on-going management of digital resources over time with a view to retaining their intellectual content, authenticity, and accessibility for a variety of uses. Additionally, for universities, these include

the protection of institutional memory, and the on-going vitality of research and scholarship. According to the Library of Congress (2010) traditional information sources such as books, photos and sculptures can easily survive for years, decades or even centuries but digital items are fragile and require special care to keep them useable. Li and Banach (2011) note that illuminated manuscripts have lasted for over 1000 years, but a CD will degrade in as little as 15 years. Rapid technological changes also affect digital preservation. As new technologies appear, older ones become obsolete, making it difficult to access older content. Chen (2001) refers to the paradox for digital preservation: On the one hand, we want to maintain digital information intact as it was created; on the other, we want to access this information dynamically and with the most advanced tools.

In Africa, there have been great strides in the implementation of technology initiatives including the management of digital records (Katuu and Ngoepe, 2015b, Ngoepe, 2015). However, there are still significant challenges in addressing long term preservation of digital information. This includes the global challenges of the rapid changes in technology leading to both hardware and software obsolescence (Duranti, 2001). Nations such as the Netherlands have been able to develop strategies to address hardware obsolescence such as emulation (Van der Hoeven et al., 2007). In Africa, South Africa is considered the most advanced in aspects of managing digital records (Kemoni, 2009). However several commentators note that the country does not have the institutional capacity to effectively management its digital content (Katuu, 2012, Ngoepe and Keakopa, 2011). In addition, there are concerns about legal and regulatory issues related to access to digital content over the long term (Katuu and Ngoepe, 2015c). In other African countries there is often irregular electricity supply leading to loss of information. In addition there are inadequately trained professionals to address the challenges of managing digital content (Katuu and Ngoepe, 2015a). Therefore, since many African institutions are struggling to manage digital content, universities and research institutions are not able to provide leadership in seeking solutions for preservation and long term access to digital content (Kanyengo, 2009).

This study therefore sought to examine how scholarly content generated and/or acquired in universities in Kenya is preserved and archived for current and future use

4. Methodology

The population of the study consisted of academic staff, postgraduate students (PhD and Master's) and university librarians of six universities in Kenya which were selected based on their relative performance in the 2013 Webometric ranking of universities. These universities were University of Nairobi, Maseno University; Kenyatta University; Jomo Kenyatta University of Agriculture and Technology (JKUAT); Strathmore University; and Egerton University. Within the six universities convenience sampling was done to obtain a sample of 350 academic staff and 370 postgraduate students. Separate self-administered survey questionnaires were designed for data collection from the postgraduate students and the academic staff. The data collected from these respondents was mostly quantitative and was analysed using SPSS to obtain descriptive and inferential statistics. Qualitative data obtained from open-ended questions in the questionnaire were analysed thematically. Generally, of the 350 and 370 copies of the questionnaires administered to academic staff and postgraduate students, 273 (78%) and 332 (89.7%) respectively were returned and were found useful for analysis. The University Librarians at the six universities were purposively selected as key informants directly involved with facilitating and managing research and scholarly

communication. Personal interviews were used to collect qualitative data from them.

5. Findings

The results of the investigation on the research question are presented in this section.

5.1 Backup and storage of research information

Respondents were required to assess their practices during and after research. The results in Table 1 indicate that respondents are conscious about long-term accessibility of their research information. As the results indicate, majority of the respondents document their research procedures, backup their information, move files to newer computers and print hard copies of files they would like to keep. The Cronbach's Alpha value for this question was 0.74 for both academic staffs and students, suggesting a high inter-item reliability.

Researcher activity	Respondent type	Backup of research information							
		Strongly Disagree		Disagree		Agree		Strongly Agree	
		FQ	%	FQ	%	FQ	%	FQ	%
I document research procedures	Academic staff	7	2.7	13	4.9	150	57.0	93	35.4
	Student	12	3.8	39	12.5	175	56.1	86	27.6
I back-up information	Academic staff	7	2.6	7	2.6	127	47.4	127	47.4
	Student	4	1.2	13	4.0	133	40.1	177	54.1
I move files to newer computers	Academic staff	9	3.4	31	11.7	115	43.2	111	41.7
	Student	33	10.2	55	17.0	150	46.3	86	26.5
I print hard copies	Academic staff	17	6.3	31	11.6	119	44.4	101	37.7
	Student	33	10.2	39	12.0	147	45.2	106	32.6
I review files in order to keep or destroy	Academic staff	10	3.7	20	7.5	145	54.1	93	34.7
	Student	17	5.3	57	17.8	141	43.9	106	33.0

Table 1: Respondent's method of backup

Key: Fq=frequency. (Percentages quoted in the text were obtained by summing up percentages in the columns of agree and strongly agree.) Academic staff (N=273) Students (N=332)
Cronbach's Alpha: Academic staff's items: 0.74; Students' items: 0.74

Source: Moseti (2015, p. 145)

5.2 Mode of preserving scholarly content

Respondents were asked to state their preferred mode of preserving scholarly content such as datasets or any other scholarly output. The study found that academic staff and students use nine modes for preserving scholarly content, shown in Table 2.

Mode of preserving scholarly content	Academic Staff			Students		
	Responses			Responses		
	N	%	% of cases	N	%	% of cases
Computer at work	150	19.1	56.2	120	14.4	36.9
University server	36	4.6	13.5	40	4.8	12.3
University digital archive	36	4.6	13.5	22	2.6	6.8
External web server	66	84.	24.7	65	7.8	20.0

	Academic Staff			Students		
Hard copy	97	12.3	36.3	156	18.7	48.0
Portable storage	173	22.0	64.8	209	25.1	64.3
Computer at home	178	22.6	66.7	193	23.1	59.4
Discipline's digital archive	35	4.4	13.1	29	3.5	8.9
My blog	16	2,0	6.0	0	0.0	0.0
Total	787	100	294.8	834	100	256.6

Table 2: Mode of preserving scholarly content

N=273 for academic staff; N=332 for students (Multiple responses possible)
Cronbach's Alpha: Academic staff's items: 0.69; Students' items: 0.71

Source: Moseti (2015, p. 146)

The results indicate that respondents used more than one mode for preserving scholarly content. On average, each academic staff used about three types (294.8/100) of content preservation modes compared to a student's two (256.6/100). The major modes of scholarly preservation were home computers, portable storage, computer at work and hard copies. The least common modes of preservation were blogs, discipline's digital archive, university's digital archive and university servers. Comparatively, slightly more academic staff (150, 19%) than students (120, 14%) preserved their scholarly content in computers at work and university's digital archives (36, 5% academic staff, 22, 3% students), (Figure 2). On the other hand, more students maintained their scholarly content in hard copies (156, 19% students; 97, 12% academic staff), and portable storage (209, 25% students; 173, 22% academic staff). Inter-item reliability as measured by the Cronbach's Alpha was relatively high (0.69 and 0.71, for academic staffs' and students' items, respectively), which showed a high internal consistency.
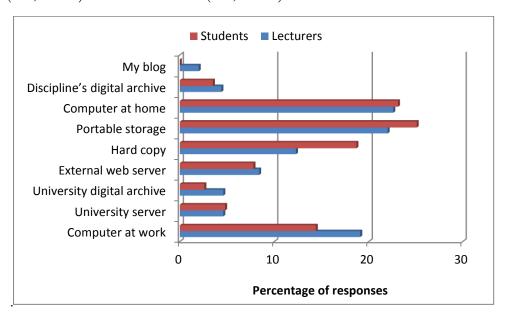

Figure 1: Respondents' mode of preserving scholarly content

A cross tabulation of the mode of preserving scholarly content and university on combined data showed that little content was preserved in university server and digital archives in all the universities (with < 10% of the respondents in any university using any of them) (Table 5). However, JKUAT (13, 7%), University of Nairobi (34, 4%) and Kenyatta (10, 4%) used more of the discipline's digital archives compared with Maseno (1, 1%), Strathmore (1, 2%), and Egerton (5, 3%) respectively. Egerton (1, 1%) and Maseno (1, 1%) also had the poorest utilisation of university servers compared to Strathmore (6, 9%), University of Nairobi (53, 7%), and Kenyatta (9, 4%) respectively. On the other hand, university's digital archives were mostly used at Strathmore (5, 8%), Kenyatta (13, 5%) and Egerton (6, 4%) but less in Maseno (2, 2%). Information from interviews with the university librarians at Strathmore and Kenyatta universities confirmed that their scholars were required to deposit all their journal articles in the institutional .

repositories. Egerton University had implemented the IR policy and was creating awareness about it among scholars. University of Nairobi was in the process of creating awareness about the IR as a vehicle for preservation and access of scholarly content among academic staff and students. Interviews with University librarians revealed that people are not comfortable with depositing material in the IRs. According to one of the librarians, 'someone will tell you: "*Sorry, you are not going to put my document there, I don't care what the policy says; you cannot have that document!*" The librarians attributed this reluctance and apathy to lack of awareness and distrust of the intentions of the IR with regard to their scholarly output. Another librarian commented: "*the only challenges we face as a library is the fact that our researchers are not aware of the importance of the IR... sometimes they complain about the fact that their work might be plagiarized.* Table 3 indicates the predominant storage formats used in the universities in the study

University	Mode of preserving data								Total
	Computer at work	University server	University digital archive	External web server	Hard copy	Portable storage	Computer at home	Digital archive of discipline	
A	140 (17.1%)	53 (6.5%)	26 (3.2%)	65 (7.9%)	135 (16.5%)	194 (23.7%)	172 (21.0%)	34 (4.2%)	819
B	25 (18.4%)	1 (0.7%)	2 (1.5%)	11 (8.1%)	27 (19.9%)	38 (27.9%)	31 (22.8%)	1 (0.7%)	136
C	40 (16.3%)	9 (3.7%)	13 (5.3%)	13 (5.3%)	38 (15.5%)	57 (23.3%)	65 (26.5%)	10 (4.1%)	245
D	21 (10.8%)	6 (3.1%)	6 (3.1%)	24 (12.4%)	26 (13.4%)	47 (24.2%)	51 (26.3%)	13 (6.7%)	194
E	11 (16.7%)	6 (9.1%)	5 (7.6%)	8 (12.1%)	9 (13.6%)	14 (21.2%)	12 (18.2%)	1 (1.5%)	66
F	33 (22.8%)	1 (0.7%)	6 (4.1%)	10 (6.9%)	18 (12.4%)	32 (22.1%)	40 (27.6%)	5 (3.4%)	145
Total	270	76	58	131	253	382	371	64	1605

Table 3: Mode of preserving scholarly content at the universities

N/B: Multiple responses possible (N=273 for academic staff; N=332 for postgraduate students) University A - University of Nairobi; University B - Maseno University; University C – Kenyatta University; University D- JKUAT; University E – Strathmore University; University F – Egerton University

Source: Moseti (2015, p. 148)

5.3 Institutional efforts in digital preservation of data

The study sought opinions from respondents on whether their institutions were making any visible efforts towards digital preservation of research data and research findings. The results are shown in Table 4.

Respondent type	University making effort	Frequency	Percentage
Academic staff	No	38	20.2
	Yes	150	79.8
	Total	188	100
Post graduate student	No	146	52.3
	Yes	133	47.7
	Total	279	100

Table 4: University's effort in data preservation

N=188 for academic staff; 279 for postgraduate students

Source: Moseti (2015, p. 150)

A Chi – square (χ^2) cross tabulation was computed to determine if interventions to preserve digital data were dependent upon the respondent's university. There was a statistically significant influence of the respondent's university on efforts towards digital preservation of research data and findings, χ^2 (5) = 29.87, p < 0.001. Results in Table 5 indicated that universities which were perceived to be making the greatest efforts to preserve digital research data were Strathmore (15, 88% of the respondents said it was), followed by JKUAT (37, 77%), Egerton (31, 76%), and Kenyatta (50, 73%). University of Nairobi and Maseno were perceived to be making the least effort, with only (124, 50%) and (26, 57%) of the respondents having the opinion that their institutions were making efforts to preserve data.

University		Respondent's university						
		A	B	C	D	E	F	Total
Visible efforts	No	122 (49.6%)	20 (43.5%)	19 (27.5%)	11 (22.9%)	2 (11.8%)	10 (24.4%)	184 (39.4%)
	Yes	124 (50.4%)	26 (56.5%)	50 (72.5%)	37 (77.1%)	15 (88.2%)	31 (75.6%)	283 (60.6%)
Total		246 (100%)	46 (100%)	69 (100%)	48 (100%)	17 (100%)	41 (100%)	467 (100%)

Table 5: Cross tabulation of respondent's university and efforts to preserve data

Key (N=467) University A - University of Nairobi; University B - Maseno University; University C – Kenyatta University; University D- JKUAT; University E – Strathmore University; University F – Egerton University
Source: Moseti (2015, p. 151)

The study required respondents to provide more detail on what their universities were doing and/or ought to do to preserve data. Most respondents (107, 34%) were of the opinion their university had repositories. However, a significant proportion of respondents suggested that universities should intensify training of users in strategies for digital preservation (79, 25%) and provide more computers and servers (57, 18%) (Table 6).

Respondent type	Frequency	Percentage
Provide more computers/servers	57	17.9
University has repositories/archives	107	33.5
Establish policy to upload data into websites	21	6.6
Offer more training/conferences	79	24.8
Encourage students to publish	36	11.3
Encourage more sharing of data/findings	19	5.9
Total	**319**	**100**

Table 6: Proposed strategies to improve digital preservation of data

N=319

Source: Moseti (2015, p. 151)

The results presented in the preceding section are discussed.

5.4 Preservation and archiving of scholarly content for current and future use

Results in Table 1 revealed that the respondents' activities with regard to preservation and archiving of scholarly content could be regarded as curative in nature. A majority of them documented their research procedures, backed up information on computers by storing multiple copies of their files in different locations, moved files from older to newer computers and also printed out hard copies of the files. The universities also required students to deposit their theses in PDF format, an open file format currently used to guarantee long term availability and portability of the document across different computer platforms.

These results suggest that both the universities and individual researchers were aware of the need to maintain the long-term accessibility of their research information and took measures to guard against its loss or inaccessibility. In requiring students to submit theses in PDF format, the universities were also complying with the National Information Standards Organization (2007, p. 37) recommendation for authors to create born-digital content in specific formats for long-term accessibility.

These results seem somewhat divergent with results from previous studies (Groenewald and Breytenbach, 2011, Lord et al., 2004, Marshall et al., 2006) which revealed that researchers and home computer users generally lacked knowledge, general awareness and consistent usage of preservation strategies and management of digital objects created on their personal computers. The current study indicates instead that computer users were increasingly aware of the need to undertake personal initiatives to ensure that their digital data and information remained consistently accessible and available for long-term use. Respondents used diverse preservation modes ranging from home and work-place computers, portable storage, and university digital archives and servers. Some respondents still relied on hard-copy

printouts for back up purposes suggesting that despite the proliferation of ICT and digital information, the traditional hard-copy format is still an important avenue of information access and preservation among Kenyan scholars. These results complement Tenopir et al, (2005) who studied the electronic journals reading patterns of astronomers who were members of the American Astronomical Society. The study revealed that even though astronomers accessed about 80% of their readings from electronic sources, they often downloaded and printed out articles on paper before reading. In the digital age, paper is no longer an attractive medium for preservation of information although its durability and usefulness compared to electronic formats continues to be debated. Wu (2005) asserts that, though paper does not suffer from the technological instability of online data, it does age. Pages become brittle; text fades; and acids, temperatures, humidity, and various environmental factors lead to deterioration of physical forms.

Although the current study results suggest increasing awareness among computer users about digital preservation, it is important that they continually be exposed to more knowledge and training to cement their knowledge and expose them to new strategies and new knowledge. As the Library of Congress (2013, p. 3) argues, "one of the still unfolding impacts of the computer age is that everyone now must be their own digital archivist".

5.5 Use of digital archives

The results revealed that university digital archives and university servers were the least popular avenues for preservation of scholarly content, favored by just 5% of both academic staff and students. A related study in Europe by (Thaesis and Van der Hoeven, 2010, p. 19) found that only 20% of the researchers submitted data to a digital archive and concluded that "researchers were not familiar with data archives and

when they were, there was still a lot of distrust in the capability of digital archives to properly handle research data". Similarly, other studies in university settings found challenges with the acceptance and use of digital archives (Davis and Connolly, 2007, Krevit and Crays, 2007, Lawal, 2002, Pelizzari, 2004, Rowlands and Nicholas, 2005, van Westrienen and Lynch, 2005). These were attributed to distrust, lack of awareness, fear of plagiarism, confusion about copyright, concerns about quality of the material in the IR, questions of who would use the material deposited and how, and the time and effort required to deposit material into digital repositories.

In the current study, interviews with university librarians revealed that scholars were not comfortable with depositing material in the IR. Some of the scholars totally refused to deposit content despite being aware of the policies governing this, saying *"I don't care what the policy says! You cannot have this document!"* This revealed that scholars seemed to have a negative attitude towards the IR and also seemed to disagree with what the policy requires about depositing of content as well as what they perceive to be the objectives and functions of the IR. As another librarian also observed *"sometimes they (the scholars) complain that their work might be plagiarized"*. The librarians attributed these attitudes to lack of awareness and distrust of the intentions of the IR and intended use of their scholarly output. The related studies cited found similar fears among the scholars they surveyed, indicating that the concerns raised by Kenyan researchers about acceptance and use of the IR are largely similar to those of scholars from other parts of the world.

In order to capture content and populate their collections, libraries in most of the universities surveyed in the current study were actively engaged in harvesting or otherwise mediating deposits of various types of scholarly content although with much difficulty. In some instances, the repositories were being populated by academic staff and students obliged to do so

through stringent measures enforced by the universities. As Swan and Carr (2008, p. 32) found out, advocacy and mediated deposit by library/repository staff can be effective tools in recruiting content, but only with the support of senior management.

It was found that respondents from Strathmore perceived their university to be making the greatest efforts in digital preservation of research data (15, 88%), followed by JKUAT (37, 77%), Egerton (31, 76%) and Kenyatta (50, 73%) respectively. Respondents in University of Nairobi and Maseno felt these universities were making the least effort in digital preservation of research data (124, 50% and 26, 57% respectively). Some respondents reported that there were no efforts by their universities to preserve data despite there being IRs in their universities. This result implies a disconnect between the IR and the users that needs to be addressed. Furthermore, the results indicated that users did not understand the functions of the IR and the libraries needed to educate users. Cochrane and Callan (2007) found that constant communication with faculty to inform and remind them of the requirements of the IR and identifying the best strategies for recruiting scholarly content into the IR (along with institutional mandate) was the best approach for achieving successful implementation of the IR. The message should be about how self-archiving is personally and professionally beneficial to scholars' research projects. Similarly, Covey (2011) advocates for aggressive marketing of the IR characterised by comprehensive campaign and targeted sale strategies. Every opportunity for contact with faculty, formal presentations at departmental meetings, articles and advertisements in campus publications must be utilised for success to be achieved.

5.6 Strategies to improve preservation

To guarantee preservation of research data, respondents of the study indicated that training on preservation strategies needed to be offered (academic staff 206, 30%;

students 226, 31%); establishment of more digital repositories (25% of both academic staff and students); more financial and operational resources (23% of both academic staff and students); and increased knowledge/expertise (academic staff 150, 22%; students 152, 21%). A related study carried out by Thaesis and Van der Hoeven (2010) on the digital preservation of research output in Europe established that general awareness on the importance of long-term preservation of research output was lacking and it needed to be created. The study recommended that in addition to research output being archived in accessible repositories, arrangements for its access in any form over the long term must be carefully considered. Institutions need to promote storage of data by creating awareness among scholars. Organisations should develop training courses to teach researchers how to manage digital data, how to work with, archive and share data sets. Further, organisations need to be aware that digital preservation is not only a technical challenge but also requires adjustments to policies and procedures for such preservation to be realised and to be beneficial to researchers.

The study revealed that the respondents were engaged in curative activities for scholarly content. The respective universities expected postgraduate students to deposit theses as PDF files into the IR. In contrast to previous studies that suggest lack of knowledge and awareness on digital preservation, the current study, indicates an increasing number of users who were aware of the importance of preserving digital information.

The results indicate that universities in Kenya are yet to fully appreciate the importance of such repositories in management of locally generated knowledge to facilitate its long-term access, use and visibility. The researcher concludes that although Kenyan scholars were aware of the importance of preservation of scholarly content, they preferred implementing preservation strategies at a personal level

rather than collectively at departmental or institutional levels. Such fragmented strategies were unlikely to contribute to the preservation of institutional memory since information would easily be lost if a person left the institution or passed on. Such strategies also limited the visibility of local content.

To guarantee preservation of research data, the study recommends more training and awareness campaigns on preservation techniques; establishment of more digital repositories and more financial and operational resources.

6. Concluding remarks

The study investigated strategies used by universities in Kenya for the preservation of their scholarly content. The study revealed that scholars in universities in Kenya are aware of the necessity for preservation of scholarly content and they utilise various strategies at personal level to enhance long-term access to this content. However, they prefer preserving the content on personal devices rather than public devices such as digital servers and repositories. Distrust and lack of awareness seem to be the key factors impeding use of institutional digital repositories for preservation of scholarly content. In line with best practices of digital preservation, the study recommends that institutions of higher learning in Kenya need to broaden their repository activities to include preservation strategy, tools and techniques in their daily management activities. This will involve integrating careful observance of accepted preservation practices throughout the lifecycle of digital information especially that which is created within the universities. All stakeholders in the information lifecycle from content creators to managers at the different levels need to commit time and resources towards digital preservation to ensure the success of these initiatives. Further, education and training of stakeholders will mitigate on the negative effects of distrust of the intentions of the IR managers and contribute to

achievement of the broader objectives of the IR.

References

Banach, M. and Li, Y. (2011), "Institutional repositories and digital preservation: Assessing current practices at research libraries", *D-Lib Magazine,* Vol. 17 No. 5, p. 1.

Becker, C., Kulovits, H., Guttenbrunner, M., Strodl, S., Rauber, A. and Hofman, H. (2009), "Systematic planning for digital preservation: evaluating potential strategies and building preservation plans", *International Journal on Digital Libraries,* Vol. 10 No. 4, pp. 133-157.

Bell, G., Hey, T. and Szalay, A. (2009), "Beyond the data deluge", *Science,* Vol. 323 No. 5919, pp. 1297-1298.

Brandt, D. S. (2007), "Librarians as partners in e-research Purdue University Libraries promote collaboration", *College & Research Libraries News,* Vol. 68 No. 6, pp. 365-396.

Brown, A., Katuu, S., Sebina, P. and Seles, A. (2009), "Module 4: Preserving Electronic Records", London: International Records Management Trust, available at: http://www.irmt.org/documents/educ_training/term%20modules/IRMT%20TERM%20Module%204.pdf (accessed 10th May 2016).

Chen, S.-S. (2001), "The paradox of digital preservation", *Computer,* Vol. 34 No. 3, pp. 24-28.

Cochrane, T. and Callan, P. (2007), "Making a difference: implementing the eprints mandate at QUT", *OCLC Systems & Services: International Digital Library Perspectives,* Vol. 23 No. 3, pp. 262-268.

Conway, P. (1996), "Preservation in the digital world", *Council on Library and Information Resources.*

Covey, D. T. (2011), "Recruiting content for the institutional repository: the barriers exceed the benefits", *Journal of Digital Information,* Vol. 12 No. 3.

Crow, R. (2002), "The case for institutional repositories: a SPARC position paper": The Scholarly Publishing & Academic Resources Coalition available at:

scholarship.utm.edu/20/1/SPARC_102 .pdf (accessed 22nd May 2016).

Davis, P. M. and Connolly, M. J. L. (2007), "Institutional repositories: evaluating the reasons for non-use of Cornell University's installation of DSpace", *D-Lib Magazine,* Vol. 13 No. 3/4.

Duranti, L. (2001), "The long-term preservation of authentic electronic records", in Apers, P. M. G., Atzeni, P., Ceri, S., Stefano Paraboschi, Ramamohanarao, K. and Snodgrass, R. T. (eds), *Proceedings of 27th International Conference on Very Large Data Bases*, Roma, Italy, Morgan Kaufmann pp. 625-628.

Groenewald, R. and Breytenbach, A. (2011), "The use of metadata and preservation methods for continuous access to digital data", *The Electronic Library,* Vol. 29 No. 2, pp. 236-248.

Hockx-Yu, H. (2006), "Digital preservation in the context of institutional repositories", *Program,* Vol. 40 No. 3, pp. 232-243.

Hoorens, S., van Dijk, L. V. and van Stolk, C. (2008), "Embracing the FutureEmbedding digital repositories in the University of London", Santa Monica, CA RAND Corporation.

Jackson, J. (2012), "Digital Longevity: the lifespan of digital files (compiled for R&D in Digital Asset Preservation)", Glasgow: University of Glasgow, available at: http://www.dpconline.org/events/previous-events/306-digital-longevity (accessed 28th March 2016).

Jantz, R. C. and Wilson, M. C. (2008), "Institutional repositories: Faculty deposits, marketing, and the reform of scholarly communication", *The Journal of Academic Librarianship,* Vol. 34 No. 3, pp. 186-195.

JISC. (2014), "Digital Preservation & Curation", available at: http://www.webarchive.org.uk/wayback/archive/20140614001302/http://www.jisc.ac.uk/whatwedo/topics/digitalpreservation.aspx (accessed 28th March 2016).

Kanyengo, C. W. (2009), "Managing digital information resources in Africa: preserving the integrity of scholarship", *The International Information & Library Review,* Vol. 41 No. 1, pp. 34-43.

Katuu, S. (2012), "Enterprise Content Management (ECM) implementation in South Africa", *Records Management Journal,* Vol. 22 No. 1, pp. 37-56.

Katuu, S. and Ngoepe, M. (2015a), "Managing Digital Heritage - An Analysis of the Education and Training Curriculum for Africa's Archives and Records Professionals", paper presented at the Digital Heritage International Congress 2015, Granada, Spain available at: http://diglib.eg.org/handle/10.2312/14622.

Katuu, S. and Ngoepe, M. (2015b), "Managing Digital Records in a South African Public Sector Institution", paper presented at the INFuture2015: e-Institutions – Openness, Accessibility, and Preservation, Zaghreb, Croatia available at: http://goo.gl/QhUjB1.

Katuu, S. and Ngoepe, M. (2015c), "Managing Digital Records within South Africa's Legislative and Regulatory Framework", in Popovsky, B. E. (ed.), *3rd International Conference on Cloud Security and Management ICCSM-2015*, Tacoma, WA, University of Washington - Tacoma pp. 59-70.

Kemoni, H. N. (2009), "Management of Electronic Records - Review of Empirical Studies from the Eastern, Southern Africa Regional Branch of the International Council on Archives (ESARBICA) region", *Records Management Journal,* Vol. 19 No. 3, pp. 190-203.

Krevit, L. and Crays, L. (2007), "Herding cats: designing DigitalCommons@ The Texas Medical Center, a multi-institutional repository", *OCLC Systems & Services: International Digital Library Perspectives,* Vol. 23 No. 2, pp. 116-124.

Lavoie, B. and Dempsey, L. (2004), "Thirteen Ways of Looking at...Digital Preservation", *D-Lib Magazine,* Vol. 10 No. 7/8.

Lawal, I. (2002), "Scholarly communication: The use and non-use of e-print archives for the dissemination of scientific information", *Issues in Science and Technology Librarianship,* Vol. 36.

Li, Y. and Banach, M. (2011), "Institutional repositories and digital preservation: Assessing current practices at research libraries", *D-Lib Magazine,* Vol. 17 No. 5.

Library of Congress [United States]. (2013), "Perspectives on Personal Digital Archiving", Washington DC: National Digital Information Infrastructure and Preservation Program: Library of Congress [United States], available at: http://www.digitalpreservation.gov/documents/ebookpdf_march18.pdf (accessed 28th March 2016).

Lord, P., Macdonald, A., Lyon, L. and Giaretta, D. (2004), "From data deluge to data curation", in *UK e-science All Hands Meeting*, Nottingham, UK, pp. 371-375.

Lynch, C. A. (2003), "Institutional repositories: essential infrastructure for scholarship in the digital age", *portal: Libraries and the Academy,* Vol. 3 No. 2, pp. 327-336.

Marshall, C. C., Bly, S. and Brun-Cottan, F. (2006), "The long term fate of our digital belongings: Toward a service model for personal archives", in *Archiving Conference*, Society for Imaging Science and Technology pp. 25-30.

Moseti, I. M. (2015), *Strategies for Managing Scholarly Content at Universities in Kenya.* Phd thesis, University of KwaZulu-Natal

National Information Standards Organization [United States]. (2007), "A framework of guidance for building good digital collections", Bethesda, MD: National Information Standards Organization, available at: http://www.niso.org/publications/rp/framework3.pdf (accessed 28th March 2016).

Ngoepe, M. (2015), "Deployment of open source Electronic Content Management Software in National Government Departments in South Africa", *Journal of Science and Technology Policy Management,* Vol. 6 No. 3, pp. 190-205.

Ngoepe, M. and Keakopa, S. M. (2011), "An Assessment of the State of National Archival and Records Systems in the ESARBICA region: A South Africa-Botswana Comparison", *Records Management Journal,* Vol. 21 No. 2, pp. 145-160.

PC.net. (2014), "File format", available at: http://pc.net/glossary/definition/file_format (accessed 01 January 2014.

Pelizzari, E. (2004), "Academic authors and open archives: A survey in the social science field", *Libri,* Vol. 54 No. 2, pp. 113-122.

Pennock, M. (2007), "Digital Curation: A life-cycle approach to managing and preserving usable digital information", *Library & Archives, January.*

Rowlands, I. and Nicholas, D. (2005), "New journal publishing models: an international survey of senior researchers": Publishers Association and International Association of STM Publishers, available at: http://www.homepages.ucl.ac.uk/~ucnciro/pa_stm_final_report.pdf (accessed 28th March 2016).

Swan, A. and Carr, L. (2008), "Institutions, their repositories and the web", *Serials Review,* Vol. 34 No. 1, pp. 31-35.

Tenopir, C., King, D. W., Boyce, P., Grayson, M. and Paulson, K. L. (2005), "Relying on electronic journals: Reading patterns of astronomers", *Journal of the American Society for Information Science and Technology,* Vol. 56 No. 8, pp. 786-802.

Thaesis and Van der Hoeven, J. (2010), "Insight into digital preservation of research output in Europe": PARSE.Insight, available at: http://www.parse-insight.eu/wp-content/plugins/download-monitor/download.php?id=19 (accessed 28th March 2016).

Van der Hoeven, J., Lohman, B. and Verdegem, R. (2007), "Emulation for digital preservation in practice: The results", *International journal of digital curation,* Vol. 2 No. 2, pp. 123-132.

van Westrienen, G. and Lynch, C. A. (2005), "Academic institutional repositories: deployment status in 13 nations as of mid 2005", *D-Lib Magazine,* Vol. 11 No. 9.

Western University [Canada]. (2016), "Digital Preservation", London, Ontario: Western Libraries, available at: https://www.lib.uwo.ca/scholarship/digitalpreservation.html (accessed 28th March 2016).

Wu, M. M. (2005), "Why print and electronic resources are essential to the academic law library", *Law Library Journal,* Vol. 97, pp. 233-256.

Biography

Irene Morara Moseti is Lecturer and Head of Department of Information Technology, School of Information Sciences, Moi University, Kenya. She holds B.Sc. Information Sciences, Masters in Business Administration and M.Phil. Information Technology from Moi University, Eldoret, Kenya. She obtained her doctoral degree in Information Studies at the University of KwaZuluNatal, South Africa. She is affiliated to the InterPares Trust as a collaborator of the InterPares Africa Team. Her research interests are in Information and Knowledge Management, Information Technologies and Systems, Open Access and Management.