Research data curation at Kenya's agricultural research institute libraries : opportunities and challenges

DOI: 10.61735/1k03qc39

[1]Emily Jeruto Ng'eno, [2]Damaris Odero, [3]Duncan Amoth

[2] https://orcid.org/0000-0001-6113-9629, [3] https://orcid.org/0000-0002-7009-7245

[1,2,3]Moi University

ABSTRACT

Research data are valuable resources that need to be curated and managed by research libraries as they are intricate and complex, irreplaceable, expensive, and time-consuming to replicate. However, The Kenya Agricultural and Livestock Research Act No.17 of 2013 does not define how research data generated by the research institutes should be managed resulting in poor mechanisms for data curation and sharing, low-quality research outputs, duplication of research, and poor re-analysis of existing research data. The purpose of this research was to examine how Kenya's agricultural research institute libraries curate their research data and propose interventions for improvement. The study investigated how Kenya's agricultural research institute libraries capture, appraise, describe, preserve, and make accessible for the reuse of its research data. The study was underpinned by the Data Curation Centre (DCC) Lifecycle Model. Six of Kenya's agricultural research institutes were purposely chosen. The study adopted a pragmatism paradigm applying a mixed methods approach and employing a survey design within a multi-case study. The target population was composed of 41 directors of institutes, heads of research, heads of IT, and librarians, and 142 researchers. Quantitative data were collected using questionnaires by simple random sampling of 124 researchers and qualitative data was collected from the remaining respondents numbering 33 using interviews sampled using census. Qualitative data was analysed thematically while quantitative data was analysed using SPSS to generate descriptive and inferential statistics. The findings revealed that the Research Data management (RDM) legal framework had not yet been enshrined in the KALRO Act (No.17 of 2013) thus affecting data curation; the Research Data Curation (RDC) policies and regulations were outdated; the institutes did not involve libraries and librarians in coordinating functions of research data curation; the librarians and researchers had inadequate knowledge and skill of handling RDC service; there was limited awareness and advocacy for RDM. The study concludes that RDC services were not adequately managed. The study recommended the establishment of a formal data governance structure to address data curation services, a legislative and policy framework for RDM and data curation, collaboration and participation among librarians and researchers, capacity-building programs, sound technical infrastructure, and incentivization of stakeholders.

(Key words**: research data; research data curation; research data management; research institutes; research libraries.**)

I. INTRODUCTION

Research data are valuable resources that need to be managed by research institutes as they are the original sources or material, be it in digital or non-digital form, that researchers create or collate while conducting research projects (Ray, 2014). Research data can be presented in three forms: raw data directly produced from a laboratory or survey; processed data that has been cleaned, refined, arranged, and combined in a manner that is useful in research; and data published in journals (Dora & Kumar, 2015). Research data are intricate, complex, irreplaceable, expensive, and time-consuming to replicate therefore, there is a need to be identified, cleansed, and transformed through Research Data Curation (RDC). RDC is defined as managing and promoting data from its point of capture, appraisal, description, preservation, access, reuse, and transformation (Atlan, 2024; Fellous-Sigrist, 2015). As such, there is a need for research institutes to be accurate and precise with their data curation these being the core functionality of Research Data Management (RDM). RDM and data curation share a multifaceted relationship but refer to different aspects of handling research data. RDC is the active and ongoing management of data through its lifecycle of interest and usefulness (Horowitz, 2019). RDM on the other hand, is a process consisting of different activities associated with RDC, security, technical capabilities, ethical

considerations, legal issues, human resource capability, and government frameworks (Ray, 2014; Whyte & Tedds, 2011). The strategic importance of RDC in agricultural research institutes and researchers (Lewis, 2010; Van den Eynden, et al, 2011) is to: encourage the improvement and validation of agricultural research, enable scrutiny of agricultural research findings, promote innovation through retrieval, comparison, and co-analysis of agricultural research data and potential new research data uses, reduce the cost of duplicating agricultural research data collection, improving accessibility and usability of curated data.

Research Data Curation (RDC) in the United Kingdom (UK), United States of America (USA), Australia, and Canada have made great advancements (Henty, 2014; Lewis, 2010; National Science Foundation (NSF), 2007). Large-scale curation of the research data emerged over forty years ago in Europe when the United Kingdom (UK) Data Archives was established to manage paper-based surveys and other data outputs. RDC in research libraries has been given impetus by the growth of digital research and growing interest in long-term preservation, curation, and storage of research data for reuse. The growth of digital research has seen the emergence of data-intensive and collaborative research leading to the establishment of The National Science and Technology Council Committee in the United States (US) and the e-Infrastructure Reflection Group in the European Union to advise on capability, capacity, and infrastructure in data curation (van den Eynden, et al., 2011). These developments have increased investment in data management (Lewis, 2010). For example, the US National Science Foundation (NSF) has invested funds and cyber-infrastructure for RDC through the DataNet program (NSF, 2007). Australia has also moved relatively speedily to develop data curation of e-research and has set up the Australian National Data Services (ANDS) (ANDS Technical Working Group, 2007).

Deventer and Piennar (2015) point out that South Africa is leading the cluster of African countries in embracing research data curation. Lötter (2014) and Fernihough (2011) affirm that The Data Intensive Research Initiative of South Africa (DIRISA) is one of the initiatives aimed at promoting RDM in the country by facilitating research data curation. In Kenya, there are attempts, albeit limited, to promote RDM, especially in the health and migration sector (Jao et al., 2015; Olum, 2013) through data curation. For example, the International Centre of Insect Physiology and Ecology (ICIPE), and the International Livestock Research Institute (ILRI) have established data management, modelling, and geo-information unit and communications and knowledge management unit respectively to ensure research outputs generated from the research activities are organized, curated, managed, and made openly accessible to wide audiences (Alila & Atieno, 2006).

Arguably, agricultural research data is the foundation of contemporary agriculture, generating valuable information and knowledge that propels economic viability, sustainability, and productivity. Its effective utilization benefits society and enhances food security and resource management. As the Government of Canada, (2016) and Mugata (2014) aptly observed, the ability to collect, describe, preserve, access, reuse, and build upon agricultural research data has become critical in advancing the agricultural knowledge base. They note that this is vital in supporting agricultural innovative solutions to economic and social challenges. It also plays a pivotal role in enhancing agricultural repositories with scientific datasets and publications to boost agriculture research, undoubtedly underscoring the need for agricultural research institutes libraries to engage in data curation and RDM. In essence, RDC in agricultural research institutes requires a legal and policy framework that responds to several RDM drivers such as ICT infrastructure, security, institutional capability, and quality assurance and compliance to enhance management, sharing, and reuse of agricultural research output (Higman & Pinfiled, 2015; Pinfield at al., 2014). In this regard, there is a need for research libraries to be proactive, accurate, and precise in guiding the collection, description, preservation, access, reuse, and sharing of research data.

This study on agricultural research data is premised on the fact that Kenya's, agricultural sector is the mainstay of the country's economy because it contributes 26% of the GDP accounts for 65% of the country's total exports, and provides more than 18% of formal and 70% of informal employment in the rural areas (Kenya, Republic of: Ministry of Agricultural, Livestock, and Fisheries, 2010; Kenya Agricultural Research Institute (KARI), 2012). The Kenya government therefore attaches great importance to the agricultural sector. Consequently, the Kenya Agricultural and Livestock Research Organization (KALRO) was set up vide the Kenya Agricultural and Livestock Research Act (No.17 of 2013) to coordinate agricultural research in the country. However, the Act does not clearly define how research data generated in the research institutes should be managed to ensure the continued preservation, long-term access, sharing and re-use of the data.

Currently, KALRO comprises sixteen (16) agricultural research institutes that house research programs in land and water management, livestock and range management, food crops, horticultural and industrial crops, and social

economics, among other areas (KALRO, 2016). Additionally, KALRO promotes sound agricultural research, technology generation, and dissemination to ensure food security through improved productivity and environment conservation (Devex, 2017). To improve RDC services, the study looks at how Kenya's Agricultural Research Institute libraries curate their research data and suggest interventions.

Problem statement and purpose of the study

Agriculture is the bastion of Kenya's economy, a major contributor to national food security, and a stimulant to employment growth. The Kenya government's strategy for revitalizing agriculture links the national research system with the agriculture sector (Kenya, Republic of: Ministry of Agricultural, Livestock, and Fisheries, 2010; Kenya, Republic of, National Development Plan, 2002-2008). KALRO was established through an Act of Parliament (The Kenya Agricultural and Livestock Research Act No.17 of 2013) to coordinate agricultural research in the country (KALRO, 2016). The Kenya Agricultural and Livestock Research Act No.17 of 2013 does not define how research data generated by the research institutes should be managed to ensure the continued preservation, long-term access, sharing, and reuse of such research data.

Concerns have been raised by Mugata (2014) who observed that though the challenges of RDC in the agriculture research institutes in Kenya are known, it remains unclear as to why they have not been addressed or assigned to their libraries. The result of this gap is visible in the poor mechanisms for data curation, sharing, and exchange, low quality of research outputs, duplication of research, high costs of gathering data, and poor re-analysis of existing research data. Additionally, a framework for capturing, organizing, and preserving data for long-term use is nonexistent. The cumulative effect of these multitudes of challenges has resulted in valuable datasets becoming lost or discarded when researchers leave or disengage with the research institutes as mechanisms for managing succession are dysfunctional. According to Beintema, (2015) and Alila and Atieno (2006), the absence or lack of enforcement of legal, policies, guidelines or otherwise of RDM to encourage researchers to deposit their research output in appropriate spaces such as institutional repositories worsens the situation.

Furthermore, a study by Wambani (2011) discovered that KALRO researchers and librarians lacked the necessary resources and training to enhance data collection, evaluation, preservation, access, sharing, and reuse. These circumstances have resulted in incomplete and inaccurate data, along with loss of research data consequently hampering access, sharing, use and reuse of research data (Ndemo, 2016). The aforementioned issues raise a fundamental question that this study attempts to addresses: How do Kenya's agricultural research institutes libraries curate their research data in terms of capturing, appraising, describing, preserving, and accessibility for reuse of its research data?

Theoretical perspective and literature review

Research Data Curation is underpinned by the Data Curation Centre (DCC) lifecycle model (Higgins, 2008). The Data Curation Centre (DCC) lifecycle model promotes a lifecycle approach to the management of digital research data to enable their successful curation and preservation from their initial conceptualization to either selection for reuse and long-term preservation or disposal (Higgins, 2008). The DCC lifecycle model has its critical starting point at the research conceptualization stage and is designed to ensure that all necessary phases of curation are planned and undertaken in the correct sequence (Pryor, 2012). According to the model, the key variables in data curation are data capture, appraisal, disposal, description, preservation, access, use and reuse, sharing, and transformation of research data. In this regard, the model outlines activities planned at different levels of granularity which includes defining roles and responsibilities; building frameworks of standards and technologies; and ensuring that processes and policies are adequately documented (Higgins, 2008). The relevancy of the DCC Lifecycle model in this study is that it lays down activities prescribed, and advocates for policy, roles, and responsibilities to be done in data curation practices which include data capture, appraisal, disposal, description, preservation, access, use, and reuse, sharing and transformation. Therefore, the DCC lifecycle model helps curators understand the processes involved in developing curation and preservation methodologies for their research institutions.

In reviewing the literature on data curation, Heidorn (2011) noted that research institutions must curate data to protect and disseminate the intellectual capital of institutions/society which is critical to the scientific and economic development of a country. Moreover, Tenopir et al., (2012) posit that researchers, librarians, IT specialists, and archivists need to be trained in data curation such as data identification and collection, selection, appraisal, use of metadata, storing datasets in repositories, and preservation. Developing an effective data curation plan requires agricultural domain-specific researchers, librarians, archivists, and IT specialists to

understand data curation requirements, practices, and procedures. The data curation lifecycle can only exist within an institutional framework with personnel with the correct skills to perform each step regularly. Related studies that have used the DCC Lifecycle Model focusing on data curation and its role in research institutions include Shakeri (2013); Lewis (2010); Palathingal, et al., (2015) among others.

Data curation is the active and ongoing management of data throughout its lifecycle of interest and usefulness to research to enable data discovery and retrieval, maintain quality, add value, and provide for reuse over time (Palmer et al., 2007). Various stakeholders including governments, funders, agricultural research institutes, and researchers are interested in data curation, and therefore, Whyte and Allard (2014) acknowledge the challenges in data curation to include a lack of legal framework, standards, or procedures to reference and define mandatory guidelines when curating data. Reported research trends in data curation (Carlson & Leiter, 2009; Kim et al., 2011) highlights important data curation aspects which include:

i. Developing interoperable standards for describing and interchanging datasets.
ii. Need for data curators to participate in the formulation of data privacy and ownership policy.
iii. Need for a workforce skilled in data curation practices; and
iv. Development of professional education standards guided by a data-driven research agenda.

In addition, Heidorn (2011); Antell et al., (2014); Karasti et al. (2006) have discussed extensively the subject of data curation using the Data Curation Centre (DCC) Lifecycle Model. These studies that have largely assumed qualitative epistemologies and descriptive approaches emphasize the need for more efforts to be directed towards understanding practices of data curation and stewardship. Heidorn (2011) advances the view that there is a large volume of data currently not being curated adequately yet data curation is of great interest to governments, funders, agricultural research institutes, and researchers. However agricultural research institutes and funding agencies are encouraged to recognize the importance of curating data to protect and disseminate the intellectual capital of society. In reviewing literature on data curation, Palathingal et al., (2015) concluded that a global trend in curating agricultural research data in the emergence of data-intensive research calls for a well-designed technical infrastructure, trained human capital, data literacy, policies, and procedures at every stage of data curation and collaboration among agricultural research institutes.

II. METHODOLOGY

The six (6) Kenya's agricultural research institutes were purposively chosen because their focus is on agricultural research and they have a long history of undertaking agricultural research in Kenya with different disciplinary areas in agricultures such as livestock, food crops, cash crops, biotechnology, tea, coffee, and genetic resources. The study adopted a pragmatism paradigm employing a mixed methods approach. This enabled the collection of quantitative data from a target population of 234 researchers and qualitative data from 41 respondents composed of directors of institutes, heads of research, heads of IT, and librarians from Kenya's agricultural research institutes. Qualitative and quantitative data were concurrently collected and integrated into data interpretation for a more comprehensive research problem analysis.

The survey design was applied within a case study allowing for data to be collected from a large population (Babbie, 2001) and elements within the research institutes examined and described comprehensively (O'Leary, 2004). The population was stratified into five (directors of institutes, heads of research, researchers, heads of IT, and librarians). Gay and Airasian (2003) suggest that in a population with fewer than 100 people or units, there is little point in sampling rather, the entire population should be surveyed. Therefore, the census was used as a sampling strategy for 41 respondents comprising directors of institutes, heads of research, heads of IT, and librarians.

A survey using Saunders et al. (2012) sampling table was used to select a sample of 142 respondents from the researcher's stratum with a population of 234. Researchers' sample size was selected based on a 95% confidence level and a 5% margin of error. A simple random sampling of researchers was taken in every agricultural research institute. Qualitative data was collected using interviews and document reviews while quantitative data was collected using questionnaires. Validity and reliability of data collection instruments were achieved through face and content validity, triangulation method, and a pilot study. Quantitative data was analysed using descriptive statistics and presented using frequency distribution tables, pie charts, and bar graphs while qualitative data was analysed thematically based on research questions and presented as a narrative. On ethical considerations, permission was sought from the gatekeepers where the study was undertaken, consent

was sought from the respondents, and the identity of the respondents was kept anonymous.

III. RESULTS AND DISCUSSION

In the current study, the questionnaires were administered to 142 researchers. The questionnaires returned were 124 from researchers representing an 87% response rate which were used to analyse data. Similarly, for the interviews, the response rate was 80% having a sample size of 41 composed of directors of institutes, heads of research, heads of IT, and librarians with a response rate of 33. These were high and acceptable response rates as recommended by McLaughlin, Bush and Zeeman (2016) and Bryman (2012).

Capturing research data

The commonly used method of capturing data was via questionnaires, interviews, audio recordings, cameras, GIS, laboratory experiments, and field experimentation. Table 1 below shows the range of formats researchers used to capture or generate their research data.

Appraisal of research data

The respondents were also asked to state the research data appraisal checklist in their institute. The results are shown in Table 2 on the following page.

Table 1: Formats for capturing research data (n=124)[1]

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Audio	9	7.3	51	41.1	26	21.0	25	20.2	13	10.5	124	100
Images	30	24.2	71	57.3	14	11.3	6	4.8	3	2.4	124	100
Spreadsheet	39	31.5	59	47.6	18	14.5	1	0.8	7	5.6	124	100
Video	26	21.0	58	46.8	21	16.9	9	7.3	10	8.1	124	100
Data-statistical	44	35.5	61	49.2	15	12.1	1	0.8	3	2.4	124	100
Database	32	25.8	74	59.7	10	8.1	6	4.8	2	1.6	124	100
Scanned document	34	27.4	59	47.6	19	15.3	7	5.6	5	4.0	124	100
Web	30	24.2	49	39.5	27	21.8	12	9.7	6	4.8	124	100
CAD[2]	16	12.9	51	41.1	28	22.6	10	8.1	19	15.3	124	100
GIS[3]	22	17.7	43	34.7	35	28.2	12	9.7	12	9.7	124	100
Data XML	12	9.7	52	41.9	39	31.5	4	3.2	17	13.7	124	100

Table 2: Appraisal of research data (n=124)[4]

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Uniqueness	28	22.6	64	51.6	10	8.1	21	16.9	1	0.8	124	100
Repeatability	36	29.0	79	63.7	7	5.6	2	1.6	0	0.0	124	100
Science/historical value	43	34.7	67	54.0	8	6.5	6	4.8	0	0.0	124	100
Complementary/added value	42	33.9	62	50.0	9	7.3	11	8.9	0	0.0	124	100

[1] ***Source:*** *Field data, 2017*
[2] ***CAD*** *– Computer-Aided Design*
[3] ***GIS*** *– Geographical Information System*
[4] ***Source:*** *Field data, 2017*

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Reuse value	50	40.3	57	46.0	17	13.7	0	0.0	0	0.0	124	100
Substantiveness	32	25.8	75	60.5	17	13.7	0	0.0	0	0.0	124	100
Access	36	29.0	59	47.6	24	19.4	3	2.4	2	1.6	124	100
Volume	25	20.2	56	45.2	25	20.2	17	13.7	1	0.8	124	100
Cost-effectiveness	22	17.7	74	59.7	18	14.5	8	6.5	2	1.6	124	100

The findings indicate that 92(74.2%) of the respondents agreed that uniqueness was one of the items considered in the appraisal of research data whereas 22(17.7%) of the respondents disagreed that they did not appraise research data using uniqueness as a checklist. Further, 115(92.7%), 110(88.7%), 104(83.9%), and 107(86.3%) of the respondents respectively stated that research data appraisal was based on repeatability, scientific/historical value, complementary/added value, and reuse value. The respondents were further asked to state the tools used to guide the appraisal of data and the findings disclosed that they preferred using the research institute's policy. However, RDM policy and appraisal and selection policy were rarely used by the research institutes perhaps due to inadequate guidance on their use. One head of research [HR6] interviewed from research institute F concurred with the researchers and stated:

> *"... Data Appraisal is done where raw data is recorded or collected in notebooks, data sheets, or in a computer then subjected to section according to its relevance to the purpose. In the selection process, researchers are supposed to follow SOPs and the institute's policy to determine what data to keep ..."*

Then, the big question is: *what is the role of the library in appraisal?*

The findings seem to suggest that the research institutes attached much value to the research data appraisal checklist and the research institute's policy as a tool used to guide the appraisal of data in their institutes thus this range of checklists speaks to the value institutes attach to data appraisal. In this regard, Whyte and Wilson (2010) and Tjalsma and Romnouts (2011) affirm that the appraisal and selection policy must fit legal requirements relating to privacy and intellectual property rights, Public Records Acts, national data policies and codes of conduct adopted by the host institution or agricultural research institute or funders. Higgins (2008) corroborates these findings by illuminating that DCC lifecycle model states that appraisal and selection of research data should adhere to documented guidance, policies, or legal requirements.

Description of research data (metadata)

Describing and documenting research data ensures that it can be found or discovered, preserved, accessed, reused, and shared in the long term. These activities can only be achieved by using appropriate administrative, descriptive, and other metadata during the time of data creation (The University of Queensland, 2017).

Methods employed include the description of the file names on their hard drives; using handwritten notes in their lab notebook after the experiments had been completed; describing the data using the column and row labels in their spread sheets; creating descriptive metadata for each dataset; and saving the descriptions with datasets on hard drive. The findings suggest that different researchers adopted different ways of describing their research datasets depending on the research project they were undertaking. There are, however, other possible explanations why results seem to reflect disparities in the way research datasets are described in various research institutes such as the lack of harmonized procedures in data description making accessibility and sharing a challenge. It begs the question: Which role did the library play in the description of research data?

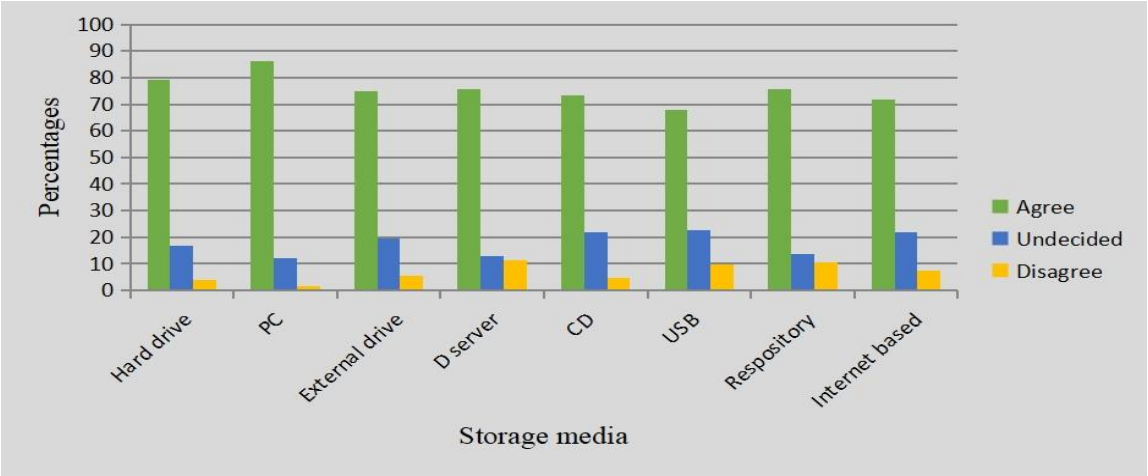
This finding resonates with that of Jones et al. (2013) who points out that at the capturing stage, file naming, versioning and structuring of files needs to be performed to ensure ease of accessing data when needed, bearing in mind both the short-term and long-term description. Agricultural research data with good metadata attached at the point of capture can expedite data sharing, mining, publishing, and citation. Similarly, the Indian Council of Agricultural Research (2014) emphasizes that research data collection should be done scrupulously and data records should be maintained through the proper use of

metadata in a durable and accessible medium that ensures safety from tampering and manipulation.

Storage and Backups of Research Data

The respondents were also asked to state the storage media of their research data awaiting preservation. The responses are presented in Figure 1 below.

Figure 1: Storage media awaiting preservation (n=124)



The study revealed that 98 (79%) of the respondents store research data in the hard drive of the instruments that generated the data. further 107 (86.3%), 93 (75%), and 94 (75.8%) of the respondents stated that the research data was stored on PC hard drives, external hard drives, and departmental servers respectively. These findings concur with interview responses where most respondents confirmed that they preferred storing their data on their laptops, external hard drives, or in cloud services and listed on research institutes' servers or repositories. The reasons given by the respondents for using these storage facilities were to maintain privacy and avoid plagiarism of their research data. The results of the findings related specifically to inadequate enforcement of data curation and RDM policy regarding standardized storage, absence of a coordinating unit (library) on storage and RDM, and absence of advocacy programs concerning standardized storage and its relevancy in data curation. In this regard, Fary and Owen (2013) assert that central IT storage, departmental storage environments, institutional repositories, and cloud-based environments are some examples of storage environments that could house the data.

The respondents were also asked to state the frequency of making backups for the research data. The results are shown in Table 3 below.

Table 3: Frequency of making backups (n=124)[5]

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Hourly	20	16.1	51	41.1	20	16.1	20	16.1	13	10.5	124	100.0
Daily	28	22.6	61	49.2	20	16.1	11	8.9	4	3.2	124	100.0
Weekly	27	21.8	58	46.8	19	15.3	13	10.5	7	5.6	124	100.0
Monthly	23	18.5	58	46.8	20	16.1	14	11.3	9	7.3	124	100.0
Annually	15	12.1	38	30.6	25	20.2	18	14.5	28	22.6	124	100.0

[5] ***Source:*** *Field data, 2017*

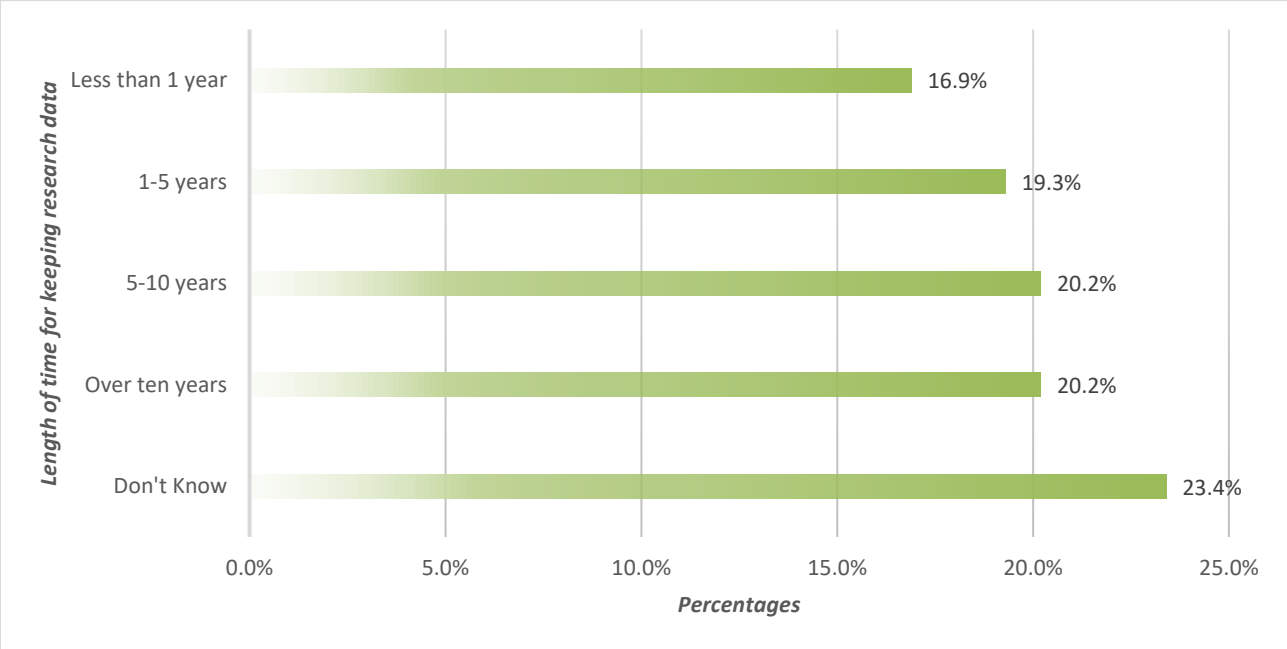
Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Never	25	20.2	45	36.3	19	15.3	16	12.9	19	15.3	124	100.0

The findings revealed that backups for the research data were done daily, weekly, and monthly in the institutes depending on individual researchers' and research institutes' policies. Most of the respondents 89 (71.8%) and 85 (68.6%) preferred backing up their data on daily and weekly respectively. Whenever a researcher makes changes to files or adds new files there is a need to backup data and it is also a good practice to backup files daily and more often when working on critical research projects. This is because research data are vulnerable to loss when researchers upgrade their computers or software. A study by Stanford Libraries (n.d) established that creating multiple backup copies of your research data is an important part of data management, but it is far less effective when all those copies are stored in the same place and not well protected.

Preservation of research data

Several preservation actions are required before research data is integrated into the curation environment to ensure that their authoritative characteristics, as identified by ISO 15489-1, can be retained for the long term (Higgins, 2012). The response of preservation strategies used is presented in Figure 2.

Figure 2: Length of time for preserving research data (n=124)



The results showed that 21(16.9%) and 24(19.3%) of respondents stated that the research data was kept for less than one year and 1-5 years respectively before it was disposed of. However, 29(23.4%) did not know the length of time for preserving the research data. These findings suggest that researchers had varying times for keeping their research data. Concerning preservation, most of the respondents interviewed lamented the lack of clear policy and guidelines governing where all research data generated should be preserved. However, each agricultural research institute had its ways of preserving data either in personal laptops, external storage, or with IT departments. Most respondents were skeptical about the preservation of the data on KALRO's server and

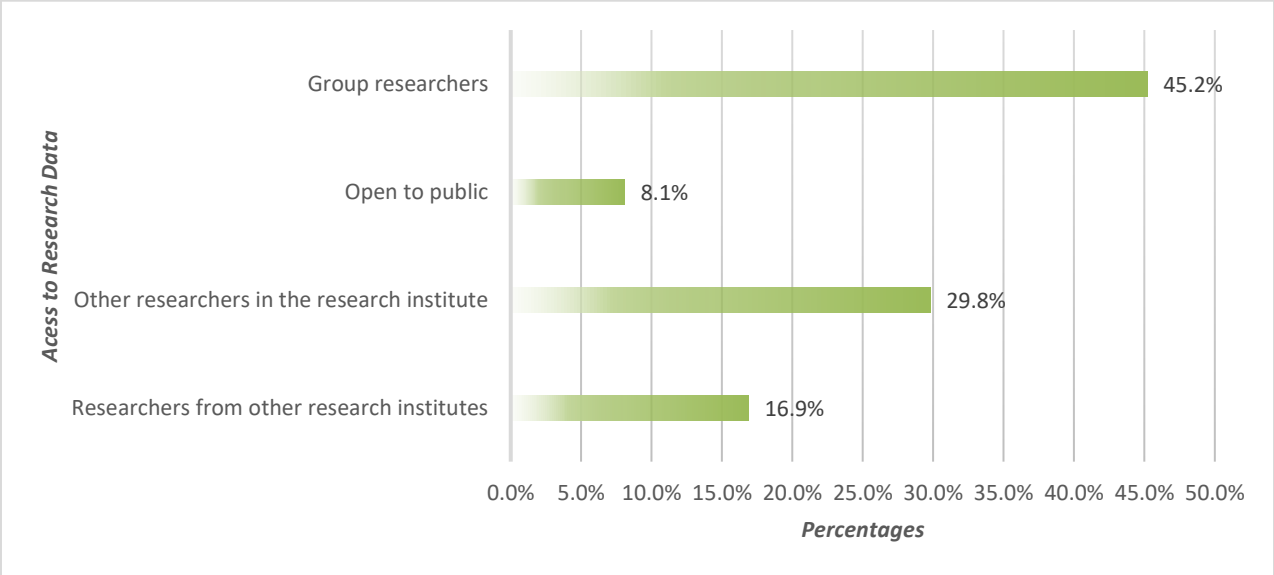
repository. This finding points to the absence of an elaborate preservation plan and a lack of awareness by researchers and librarians on what RDC and preservation policies should stipulate.

Findings showed that preservation was taking place majorly on an individual level and minimally at an institutional level. Ironically, the library and IT departments played insignificant roles in RDC hence missing the expertise of the library and IT department suggesting that preservation policy, where it existed was either ignored by researchers or was not enforced. These findings support the idea of the University of Pretoria (2007) policy for the preservation and retention of research data which states that research data belongs to the research institution and must be stored for a minimum period of ten years after the completion of the research project but if intellectual property or contractual requirements are involved then the reservation period may be altered. In addition, researchers, researchers in consultation with their researcher supervisor, and the funder's requirements should be the determinants of research data worth preserving as well as the time limit for preservation as guided by the preservation plan (Higgins, 2012). The expertise of the library, IT department, and archivists should be invoked to support and create awareness among researchers.

Research data access

The study also sought to determine the persons who were allowed to access the research data. Figure 3 below shows the results.

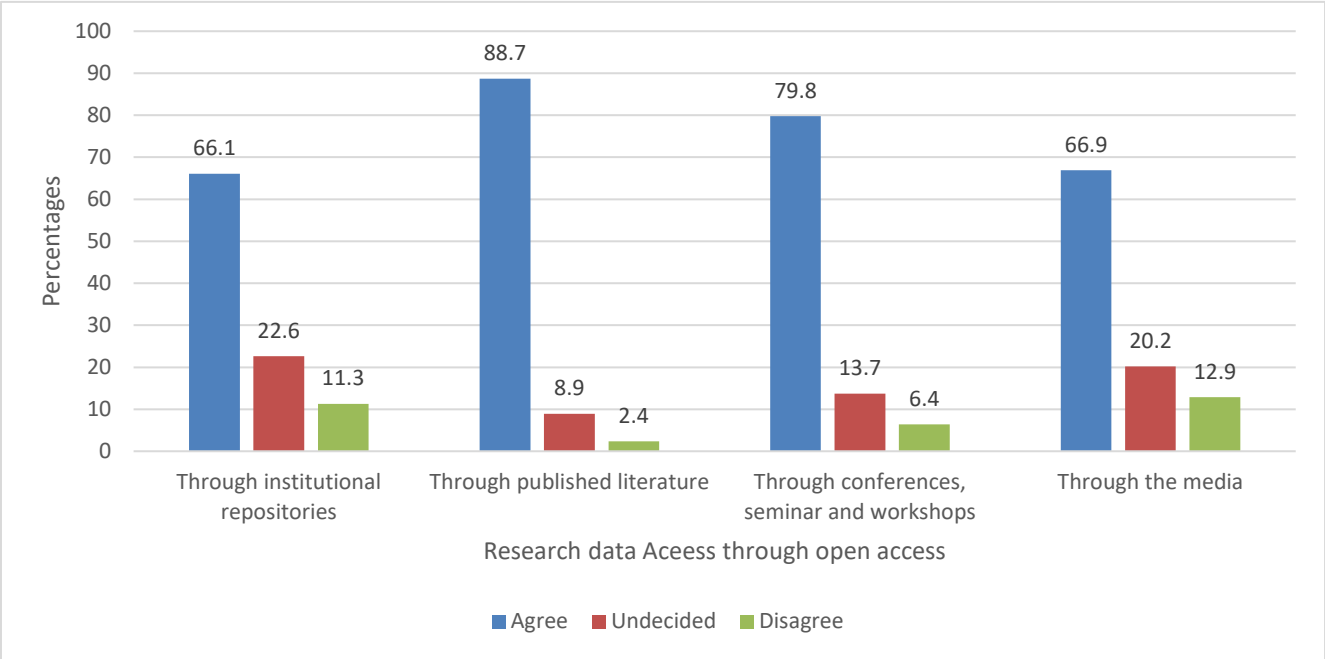
Figure 3: Access to research data (n=124)



As indicated by Figure 3, 56(45.2%) respondents indicated that group researchers were allowed to access research data. The findings also revealed that there was minimal access to research data by the public. The findings seemed to suggest that there were inadequate mechanisms to facilitate access to data. In support of the findings, the European Commission (2016) posits that fuller and wider access to research data is important as it helps to build on previous research results, encourage collaborations and help to avoid duplication of effort, speed up innovation, and involve citizens and society in the scientific process. In essence, access to research data increases the returns from public investment in research projects; reinforces open scientific inquiry, encourages diversity of studies and opinion, promotes new areas of work, and enables the exploration of topics not envisioned by the initial investigator. However, Access to Kenya's agricultural research data faces various challenges as enumerated by Muinde and Gorman (2009) and these include social-cultural (non-visionary leadership), lack of ICT infrastructure, legal, policy, and institutional frameworks, and capacity-building programs.

The respondents were also asked to state how they availed research data through open access. The results are shown in Figure 4 on the following page.

Figure 4: Research data access through open access (n=124)



On how the respondents availed research data, 82 (66.1%) of the respondents indicated that they avail research data through institutional repositories whereas 110 (88.7%) availed the data through publishing literature, and 99 (79.8%) availed the research data through conferences, seminars, and workshops. Institutional repositories and media were averagely used because the researchers were skeptical about the privacy and plagiarism of their research data. Similar concerns were raised by Kedemi (2017) about the implementation of KAINet, an institutional/national repository with scientific publications on agriculture and forestry. Kedemi identified several challenges faced by KAINet which included: the absence of institutional policies that support open access, the low awareness of copyright issues, the absence of appropriative information, management skills, system incompatibility, shortage of technical ICT skills, and collaborations with stakeholders. Besides, similar challenges are facing Kenya agricultural research institutes' repositories.

Sharing research data

Data sharing has incredible potential to strengthen agricultural research. The study wanted to investigate the method of sharing all or part of research data. Table 4 shows the methods used to share all or part of the research data.

Table 4: Methods of sharing all or part of research data (n=124)

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Collaborative web space (wiki, blog, Google Docs)	25	20.2	49	39.5	23	18.5	14	11.3	13	10.5	124	100
Data portal or database driven web site	29	23.4	47	37.9	33	26.6	8	6.5	7	5.6	124	100
Deposit them with a specialist data centre	24	19.4	50	40.3	27	21.8	15	12.1	8	6.5	124	100
Depositing them in an institutional repository	34	27.4	43	34.7	31	25.0	9	7.3	7	5.6	124	100

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Submitting them to a journal to support publication	35	28.2	42	33.9	40	32.3	1	0.8	6	4.8	124	100
E-mail	37	29.8	55	44.4	20	16.1	6	4.8	6	4.8	124	100
External storage device (USB drive, CD/DVD)	50	40.3	50	40.3	13	10.5	3	2.4	8	6.5	124	100
Hard copy or print	50	40.3	44	35.5	12	9.7	9	7.3	9	7.3	124	100
Don't share data	46	37.1	43	34.7	15	12.1	9	7.3	11	8.9	124	100

The findings indicate that 100 (80.6%) of the respondents shared their research data through an external storage device while 94 (75.8%) shared their research data through hard copy or print. Furthermore 89 (71.8%) of the respondents stated that they didn't share data and 20 (16.2%) disagreed that they did share research data. The findings show that researchers in agricultural research institutes shared research data using different methods. The preferred mediums were external storage devices, hard copy/print, e-mail, institutional repository, and journal publication. However, collaborative web space, data portal/database driven website, and depositing them with a specialized data center were methods that were partially used in sharing research data. The findings agree with Van den Eynden et al., (2011) who found that most research results were shared through specialist data centers, data archives or data banks; journal publications; institutional repositories; institutional websites; and informal sharing between researchers on a peer-to-peer basis. Van den Eynden et al maintain that sharing policies, embracing new sharing technologies and training researchers are key issues in sharing research data. Warren (2016) acknowledges that the benefits of data sharing are attained when researchers have access to complete datasets and are thus able to answer new questions, explore different lines of analysis, and more efficiently, conduct large-scale analyses in research.

Reusing research data

The respondents were asked to state the reasons for using and reusing research data. The results are shown in Table 5 below.

Table 5: Reasons for using and reusing data (n=124)[6]

Statement	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
Avoid duplication	52	41.9	49	39.5	11	8.9	2	1.6	10	8.1	124	100
Reduce the cost of doing research	55	44.4	43	34.7	16	12.9	1	0.8	9	7.3	124	100
Re-analysis of data can lead to powerful insights	57	46.0	56	45.2	3	2.4	5	4.0	3	2.4	124	100
Encourages scientific enquiry	63	50.8	51	41.1	4	3.2	2	1.6	4	3.2	124	100
Promote innovations	52	41.9	60	48.4	6	4.8	5	4.0	1	0.8	124	100
Potential new data uses	59	47.6	57	46.0	4	3.2	3	2.4	1	0.8	124	100

[6] ***Source:*** *Field data, 2017*

Concerning the use and reuse of research data, most of the respondents stated that research data is used and reused to promote potential new data uses and encourage scientific inquiry. According to the respondents, the re-analysis of data leads to powerful insights, promotes innovations, and helps avoid duplication. One head of research [HR 9] summarizes the justification for the use and re-use of research data as:

> *"… Data is always reused especially in agricultural research institutes. Sometimes when doing a report on a certain research project, one may reuse the data if results are not tallying as expected. A case in point is the scenario where data sheets done by National Gene bank of Kenya in 1988 are still used and reused until today for reference. Also reuse of coffee research data has gone a long way in improving variety of coffee and even bringing in new varieties that are manageable to the farmers like the new variety called Ruiru coffee. Food crops have equally improved their varieties due to the reuse of data. Re-using of data allows analysis of new areas of research ..."*

These findings corroborate those of Australian National Data Services (n.d) in a study on data reuse which reiterated the reasons for enabling reuse of data to be: encouraging scientific enquiry and debate, increasing the impact and visibility of research, providing great resources for education and training, and leading to new collaborations between data users and data creators. Lewis (2010) agrees that the reuse of research data assists in addressing emerging issues, supporting re-analysis of existing data or comparisons with new data in order to come up with new research themes and powerful insights which are imperative to RDM in agricultural research institutes. Research data can be reused in increasingly diverse ways and have potential beyond the original scope of a research project.

Responsibility for RDC in Kenya's agricultural research institutes is another aspect that the study focused on. The findings indicated that there was no clear unit/department/person responsible for research data curation. Both qualitative and quantitative results established that IT staff, librarian, collaborative responsibility/research group, and external research partners were responsible for RDC in agricultural research institutes. The findings suggest that Kenya's agricultural research institutes have no clear policy regarding the responsibility of research data curation. This could explain why RDC remained underdeveloped, underutilized, and without budget and resource allocation. Erway (2013) and Tenopir et al., (2014) notes that the library is well situated to be a key player and provide a range of service in RDM that include data curation, access, sharing of data literacy to researchers, institutional repositories given its extensive experience.

IV. CONCLUSION

The findings support that research data capture, appraisal, description, preservation, accessibility, reuse, and sharing in the agricultural research institute libraries were the core functions, activities, and drivers of RDC culminating in RDM. However, these core functions did not have a unit in the library or department or person responsible for coordinating the activities and ensuring quality RDC was attained. Further, the findings revealed that RDC in the research institutes was not professionally managed from creation to sharing as envisaged by the DCC lifecycle model due to the absence of a mandate assigning institute libraries to curate research data and handle RDM. The findings pointed out the need for a coherent RDM legal framework to reinforce the establishment of RDM as a statutory responsibility thus embedding data curation. The gaps and inadequacies are pertinent in areas such as librarian's knowledge and skill in curating data and handling RDM service, data literacy and advocacy; formulation and implementation of up-to-date RDM policies; sound ICT infrastructure, adopting standardized metadata, among others. KALRO also did not involve libraries and IT department in coordinating functions of RDM of which is a great setback to data curation; Overall, the findings revealed that libraries and librarians were not responsible for RDC data curation and RDM functions, activities and services hence denying maximum utilization of research data service.

V. RECOMMENDATIONS

A comprehensive and articulates data governance should be enacted in agricultural research institutes: A formal data governance structure is recommended to address the wide variety of data issues, especially research data curation, and assigning RDM responsibility to libraries and librarians in agricultural research institutes. A robust data governance framework will provide the structure and institutional oversight necessary to establish a culture of data fluency across the institutes. KALRO or specifically Research institutes bears the responsibility of enacting and implementing data governance in their institutions.

Enacting legal framework to guide on RDM: The study strongly recommends that the government should enact legislation to give research institutes a statutory mandate to capture, appraise, describe, preserve, access, reuse, and share, intellectual property rights, and data ownership among others in all research institutes. Kenya's government and KALRO should bear the responsibility

and give a mandate to institute libraries and IT departments to implement.

Revision and implementation of RDM policy and regulations: It is recommended that Kenya's agricultural research institutes should consider revising RDM policies and regulations to include current trends in RDM; mandate Kenya's agricultural research institute libraries to curate data and mechanisms of enforcing policies; administering penalties; functions, activities, service and responsibilities among other. KALRO and specifically Research institutes bears responsibility or revising and implementing the RDM policy by assigning responsibility to RDM unit or library.

Establishment of RDM department in the Library: the study recommends the establishment of RDM department in the library to oversee the functions, activities, roles, services, responsibilities and coordination of research data capture, appraisal, description, preservation, accessibility, reuse, sharing, laying down strategies and mechanisms for RDM in all agricultural research institutes; Including implementation and reinforcement of RDM policies and regulation; running advocacy campaigns and creating, awareness about RDM. Besides, library and IT departments should offer expertise in data curation, reuse, sharing, and development of institutional repositories and data management portals. KALRO bears responsibility and more specifically assign responsibility to the library.

Upskill and reskill human resource capability for RDM: It is recommended that agricultural research institutes should recruit or train librarians to possess knowledge and skills for data curation and RDM. KALRO bears responsibility for organizing seminars, workshops, or short training for upskilling and reskilling.

Technical infrastructure preparedness: Agricultural research institutes should adopt ICTs for data curation and RDM to enhance the security, accessibility, efficiency, reliability and responsiveness of RDM. ICTs should facilitate all functions, roles, activities and services of data curation and RDM. KALRO bears responsibility and more specifically IT department.

ACKNOWLEDGEMENT

The Authors' gratitude and appreciation goes to the Kenya Agricultural and Livestock Research Organization (KALRO) for granting the opportunity to undertake research in the organization. Gratitude also goes to the respondents who provided data by filling the questionnaires and agreeing to be interviewed. In addition, the article is a portion of a write-up from a PhD thesis Titled 'Research Data Management in Kenya's Agricultural Research Institutes'- University of KwaZulu-Natal, 2018 by co-author Emily Jeruto Ng'eno.

REFERENCES

Alila, P. & Atieno, R. (2006). Institute of Development Studies Agricultural Policy in Kenya: Issues and Processes. In *A Paper for the Future Agricultures Consortium Workshop, Institute of Development Studies, 20-22 March 2006*. Retrieved April 25, 2016 from http://www.fao.org/fileadmin/user_upload/fsn/docs/Ag_policy_Kenya.pdf

Antell, K., Foote, J.B., Turner, J. and Shults, B. (2014). Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions. Retrieved April 20, 2016, from http://crl.acrl.org/content/75/4/557.full.pdf+html

Atlan. (2024). Data Curation: Definition, Importance & Examples. Retrieved September 3, 2024, from https://atlan.com/know/data-curation-101/

Australian National Data Service. (n.d). Data Re-use. Retrieved October 27, 2016, from http://www.ands.org.au/working-with-data/enabling-data-re-use/data-re-use

Austrian National Data Services (ANDS) Technical Working Group. (2007). Towards the Australian Data Commons: a Proposal for an Australia National Data Service. Department of Education, Science and Training. Retrieved May 26, 2016, from http://www.ands.org.au/__data/assets/pdf_file/0009/386982/ardc_eif_annual_report_2011-12.pdf

Babbie, E. R. (2001). *The Practice of Social Research*. (9th ed.). Belmont, CA (USA): Wadsworth/Thomson Learning.

Beintema, N. (2015). African States should Double Agriculture Research Spending. Standard Newspaper (Standard digital. Retrieved April 25, 2016 from http://www.standardmedia.co.ke/mobile/article/2000142775

Bracke, M.S. (2011). Emerging Data Curation Roles for Librarians: A Case Study of Agricultural Data. *Journal of Agricultural and Food Information, 12 (1),* 65-74.

Bryman, A. (2012). *Social Research Methods*. 5th (ed.). London: Oxford University Press.

Carlson, J. & Leiter, D. (2009). Addressing Researcher's needs through the Data Curation Profile. *In Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '09)*, 365-366.

Creswell, J. W. & Plano Clark, V.L. (2011). *Designing and Conducting Mixed Methods Research*. 2nd ed. Los Angeles: Sage.

Deventer, M. V. & Pienaar, H. (2015). Research Data Management in a Developing Country: A Personal Journey. *International Journal of Digital Curation, 10(2),* 33-47.

Devex (2017). Kenya Agricultural and Livestock Research Organization (KALRO). Retrieved March 29, 2017, from https://www.devex.com/organizations/kenya-agricultural-and-livestock-research-organization-kalro-22923

Dora, M. & Kumar, H. A. (2015). Managing Research Data in Academic Institutions: Role of Libraries, 2015. In CALIBER, 2015, Shimla, March 12-+14, 2015 [Conference paper]. Retrieved September 18, 2017, fromhttp://eprints.rclis.org/24911/1/50.pdf

Erway, R. (2013). Starting the Conversation: University-Wide Research Data Management Policy. Dublin, Ohio: OCLC Research. Retrieved May 1, 2016, from http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf

European Commission. (2016). Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020: Version 2.1. Retrieved April 20, 2016, from http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Fary, M and Owen, K. (2013). Developing an Institutional Research Data Management Plan Service. Retrieved March 10, 2016, from http https://net.educause.edu/ir/library/pdf/acti1301.pdf

Fellous-Sigrist, M. (2015). What are Research Data? Retrieved March 1, 2017, from https://www.ucl.ac.uk/tobedeleted-sites/dis-old-June16/research/RDM_presentation.pdf

Fernihough, S. (2011). E-Research: An Implementation Framework for South African Organization (Unpublished Master's Research Report, University of South Africa) Retrieved May 26, 2016, from http://hdl.handle.net/10500/4474

Gay, L.R. & Airasian, P. (2003). *Educational Research: Competencies for Analysis and Application.* 7th (ed.). Upper Saddle River, NJ: Prentice Hall.

Government of Canada. (2016). Tri-Agency Statement of Principles on Digital Data Management. Retrieved March 23, 2017, from http://www.science.gc.ca

Heidorn, P.B. (2011). The Emerging Role of Libraries in Data Curation and E-Science. *Journal of Library Administration, 51: (7-8),* 662-672.

Henty, M. (2014). Dreaming of Data: The Library's Role in Supporting E-Research and Data Management. Retrieved January 3, 2016, from https://digitalcollections.anu.edu.au/bitstream/1885/47617/5/henty_alia_08.pdf

Higgins, S. (2008). The DCC Curation Lifecycle Model. *The International Journal of Digital Curation*, 3(1). Retrieved March 24, 2016, from http://www.ijdc.net/index.php/ijdc/article/viewfFile/69/48

Higgins, S. (2012). The Lifecycle of Data Management. In G. Pryor. (Ed.). *Managing Research Data*. London: Facet Publishing.

Higman, R. & Pinfield, S. (2015). Research Data Management and Openness: The Role of Data Sharing in Developing Institutional Policies and Practices. *Program: Electronic Library and Information System, 49 (4),* 364-381.

Horowitz, B.M. (2019). Data curation. Retrieved April 24, 2024, from https://www.sciencedirect.com/topics/computer-science/data-curation

Indian Council of Agricultural Research (ICAR), (2014). ICAR Guidelines for Internal Evaluation and Forwarding Research Papers to Scientific Journals and Data Management in ICRAR Institutes. Retrieved June 5, 2017, from http://www.icar.org.in/files/ICAR-Guidelines-Research-Papers-2014.pdf

Jao, I.., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., Molyneux, S. & Marsh, V.
(2015). Research Stakeholders' View on Benefits and Challenges for Public Health Research Data Sharing in Kenya: The Importance of Trust and Social Relations. Retrieved April 20, 2016, from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0135545

Jones, S., Guy, M. and Pickton, M. (2013). *Research Data Management for Librarians*. London: Data Curation Centre.

Karasti, H., Baker, K. and Hakola, E. (2006). Enriching the Notion of Data Curation in E-science: Data Managing and Information Infrastructure in the Long-Term Ecological Research (LTER) Network. Retrieved May 3, 2016, from http://www.worldpece.org/sites/default/files/57-227-1-PB.pdf

Kedemi, R. (2017). Overview of 'A KARI and KAINet Institutional Repository Development: Case Study. Retrieved June 15, 2017, from http://aims.fao.org/capacity-development/webinars/ overview-kari-and-kainet-institutional-repository-development-case

Kenya Agricultural and Livestock Research Organization (KALRO). (2016). *Service Charter (June 2016)*. Nairobi, Kenya: KALRO.

Kenya Agricultural Research Institute (KARI). (2012). *Policy Responses to Food Crisis in Kenya*. USA, Washington: Food Security Portal.

Kenya, Republic of. (2002). *National Development Plan 2002-2008*. Nairobi: Government Printers.

Kenya, Republic of. (2013). Kenya Agricultural and Livestock Research Institute (KALRO) ACT NO 17 OF 2013. Retrieved June 13, 2017, from https://www.stak.or.ke/wp-content/uploads/2016/08/The-Kenya-Agricultural-and-Livestock-Research-Act-2013.pdf.

Kenya, Republic of: Ministry of Agriculture, Livestock and Fisheries. (2010). *Agricultural Sector Development Strategy 2010-2020*. Nairobi: Government Printers.

Kim, Y., Addom, B. & Stanton, J. (2011). Education for eScience Professionals: Integrating Data Curation and Cyberinfrastructure, *The International Journal of Digital Curation, 6,* 125-138.

Lewis, M.J. (2010). Libraries and the Management of Research Data. In: *Envisioning Future Academic Library Services*. Facet Publishing: London. 145-168.

Lötter, L. (2014). *Reflections on the RDM Position in South Africa: Development Research Data Management Services*. Cape Town, South Africa: Research Methodology Centre.

McLaughlin, J.E., Bush, A.A., and Zeeman, J.M. (2016). Mixed Methods: Expanding Research Methodologies in Pharmacy Education. Retrieved June 16, 2017, from http://dx.doi.org/10.1016/j.cptl.2016.06.015

Mugata, R. (2014). KAINeT Initiative Opening Up Access to Agricultural Research Information in Kenya. Retrieved April 22, 2016, from http://www.ciard.info/news-and-events/blog/kainet-initiative-opening-access-agricultural-research-information-kenya.

Muinde, F.N.N. & Gorman, G.E. (2009). Barriers to Open Access to Scientific Information in Kenya

with Particular Reference to Agricultural Information. Retrieval August 26, 2016, from http://www.ifla.org/annual-conference/ifla75/index.htm.

National Science Foundation (NSF). (2007). Cyberinfrastructure Vision for 21st Century Discovery. Retrieved July 6, 2016, from http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf

Ndemo, B. (2016). Wheat should be Central to Kenya's food Security Strategy. The Daily Nation (digital). Retrieved April 25, 2016, from http://www.nation.co.ke/oped/blogs/dot9/ndemo/-/2274486/2986916/-/bporh/-/index.html.

Olum, G. H. (2013). *Enhancing Migration Data management in Kenya: Assessment and Recommendations.* Nairobi: ACP Observatory on Migration.

O'Leary, Z. (2004). *The Essential Guide to Doing Research*. London: Sage.

Palathingal, L., Dascalu, S., Harris, F.C. & Varol, Y. (2015). A Brief of Data Curation Literature. Retrieved May 20, 2016, from https://www.cse.unr.edu/~fredh/papers/conf/143-absodcl/paper.pdf.

Palmer, C.L., Cragin, M.H., Heidorn, P.B. & Smith, L.C. (2007). Data Curation for the Long Tail of Science: The Case of Environment Studies. Retrieved March 28, 2016, from http://www.researchgate.net/publication/228656107.

Pinfield, S., Cox, A.M. & Smith, J. (2014). Research Data Management and Libraries: Relationship, Activities, Drivers and Influences, *PLoS ONE, 9(12)*. Retrieved March 25, 2016, from http://eprints.whiterose.ac.uk/82499/1/PLOS%20ONE% 20final%20published% 20version.pdf.

Pratt, M. (2024). Data curation. Retrieved April 23, 2014, from https://www.techtarget.com/searchbusinessanalytics/definition/data-curation

Pryor, G. (2012). Why Manage Research Data? In G. Pryor. (Ed.). *Managing Research Data*. London: Facet Publishing.

Ray, J. M. (2014). Introduction to Research Data Management. In. J. M. Ray. (Ed). *Research Data Management: Practical Strategies for Information Professionals.* West Lafayette, Indiana: Purdue University Press.

Saunders, M., Lewis, P. & Thornhill, A. (2012). *Research methods for Business Students.* 6th (ed.). Edinburgh Gate, England: Pearson Education Limited.

Shakeri, S. (2013). Data Curation Perspectives and Practices of Researchers at Kent State University's Liquid Crystal Institute: A case Study (Masters Dissertation). Retrieved April 28, 2026, from http://lod-lam.slis.kent.edu/documents/StudentResearch/ShadiShakeri/Shakeri_Thesis_Corrected_Nov26_PDF-A.pdf

Stanford Libraries. (n.d). Case Study: Data Storage and Backup. Retrieved June 27, 2017, from https://library.stanford.edu/research/data-management-services/case-studies/case-study-data-storage-and-backup

Tenopir, C., Birch, B., and Allard, S. (2012). Academic Libraries and Research Data Services: Practices and Plans for the Future. *Association of College and Research Libraries*. Retrieved April 1, 2016, from http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf

Tenopir, C., Sandusky, R. J., Allard, S. and Birch, B. (2014). Research Data Management Services in Academic Research Libraries and Perceptions of Librarians. *Library and Information Science Research, 36(2),* 84-90.

Tjalsma, H. and Rombouts, J. (2011). Selection of Research Data: Guidelines for Appraising and Selecting Research data. Retrieved June 2, 2017, from file:///C:/Users/216056748/Downloads/DANSselectionofresearchdata% 20(1).pdf

United Nations Environmental Programme (UNEP). (2015). *Green Economy Sector Study on Agriculture in Kenya.* Nairobi-Kenya: UNEP. Retrieved October 3, 2016, from http://drustage.unep.org/greeneconomy/sites/unep.org.greeneconomy/files/publications/kenya_agriculture_final.pdf

University of Pretoria. (2007). Policy for the Preservation and Retention of Research Data. Retrieved July 2, 2017, from Available at: http://webcache.googleusercontent.com/search?q= cache:igaTSFfrAOEJ:www.up.ac.za/media/shared/Legacy/sitefiles/file/40/policyforthepreservationandretentionofresearchdata.doc+andcd=1andhl=enandct=clnkandgl=za

The University of Queensland. (2017). Research Data Management: Data Description. Retrieved June 9, 2017, from http://guides.library.uq.edu.au/research-data-management/data-description.

Van den Eynden, V., Corti, L., Woollard, M., Bishop, L. & Horton, L. (2011). *Managing and Sharing Data best Practices for Researchers.* London, UK: Data Archives.

Wambani, H. N. (2011). Improving Scientific Research at Kenya Agricultural Research Institute (KARI)-Kakamega Research Centre: A Research Methods Approach. (Masters Dissertation). Retrieved April 25, 2016, from http://ruforum.org/sites/default/files/Hildah%20Nyende%20Wambani%20thesis.pdf.

Warren, J. (2016). Strengthening Research through Data Sharing. *The New England Journal of Medicine, 375*, 401-403.

Whyte, A. & Allard, S. (2014). How to Discover Research Data Management Service Requirements. DCC How-to Guides. Edinburgh: Digital Curation Centre. Retrieved July 7, 2016, from http://www.dcc.ac.uk/how-

Whyte, A. & Tedds, J. (2011). Making the Case for Research Data Management, DCC Briefing Papers. Edinburgh: Digital Curation Centre. Retrieved September 23, 2016, from http://www.dcc.ac.uk/resources/briefing-papers/making-case-rdm.

Whyte, A. and Wilson, A. (2010). How to Appraise and Select Research Data for Curation. Edinburgh: Digital Curation Centre. [Online]. Retrieved June 23, 2016, from http://www.dcc.ac.uk/resources/how-guides/appraise-select-data