DATA MANAGEMENT IN RESEARCH PRACTICES IN HIGHER EDUCATIONAL INSTITUTIONS: ISSUES IN THE RESEARCH SPHERE

[1] Idika, Delight Omoji; [2] Eyong, Emmanuel Ikpi; [3]Okeke, Stella Uchechukwu; [4] Egbai, Julius Michael;[5]Otu, Benard Diwa; [6]Ojini, Richard Ayuh; [7]Anakwue, Laetitia Anthonia; [8]Shogbesan, Yusuf Olayinka; [9]Inah, Lovina Idoko.

[1 &7]Institute of Education [2]Department of Educational Foundations and Childhood Education[3]Department of Social Science Education [4,5&6]Department of Educational Foundations [8]Al-Hikmah University Ilorin [9]Department of Science Education

[1,3,4,5,6,7&9] University of Calabar, [2]Cross River University of Technology, [8]Al-Hikmah University Ilorin

EMAIL: delightoidika@yahoo.com, https://orcid.org/0000-0001-9309-7929

ABSTRACT

In the research sphere, data management as an aspect of research management has assumed tremendous significance, particularly with an expansion in the educational research enterprise. Data management represented by the acronym RDM describes the management and organization, storage, preservation, and sharing of data collected and used in a research undertaking. To this end, the study focused on data management in research practices in higher educational institutions: issues in the research sphere. Three research questions and hypotheses were formulated to guide the study, and the design adopted was ex-post facto with an accessible population of 790 data analysts in selected higher educational institutions in Nigeria. The study adopted the multistage sampling technique with a sample size of 385 (50%). The instrument used for data collection was a researchers' validated instrument titled "Data Collection Scale for Researchers (DCSR)" which was validated using the inter-rater reliability method by experts in the field of Education and the coefficient of internal consistency ranged from 0.74 to 0.79. The data were analyzed using descriptive statistics (mean, standard deviation, and bar charts). The findings of the study revealed that most data analyzed were cooked (not real) irrespective of the fact that the results were quite outstanding. Also, most data were analyzed by quacks who do not have the required skills in research practices. The study, therefore, recommended the need for compulsory participation in workshops, seminars, and conferences by all lecturers, students, and other education researchers in the academic environment to produce a worthwhile output that will inform policy decisions. Conclusively, the study proposed the need for research and data

16869

management centers along with research ethics committees to be established by institutions as these will help to checkmate and promote quality data management by research scholars.

Keywords: Data Management, Research Practices, Issues & Research Sphere.

INTRODUCTION

The role of quality data management cannot be undermined at all levels of education (primary, secondary, and tertiary). This is because the world has become a global village where every source of information has been digitalized. Data-driven technological management has been a thing of great concern to all and sundry in the academic field. This is because research has been a major area of focus for everyone in academia irrespective of one's area of specialty. This is driven by the quest for "publish or perish syndrome'' that has made most academics strive for promotion in their respective area of discipline. To attain their academic height, data management must be practiced by all institutions in Nigeria. As Idika, Joshua, and Umoinyang (2017) noted that if lecturers (scholars) lack the expertise to make the choice of data collection or analysis approach, or to carry out the actual analysis from screening and coding instruments up to computer output analysis, or the skills to categorize, order, manipulate and summarize data, it may be difficult to obtain an answer to research questions and results of hypothesis testing. Undoubtedly, this may be a reason why the findings of some researchers may not have been used in the research sphere where they could relevantly influence policies and programmes of activities. It is believed that the findings of this study will help to bridge the gap between research evidence and government policies in education.

Data management refers to the accurate and deliberate act of collecting, analyzing, using, and preserving data to make decisions, reach conclusions, create new knowledge, add to or modify existing knowledge with new evidence, and provide a platform for easy retrieval and reuse in the future. It is the process whereby data are collected, sorted, coded, prepared, analyzed (using appropriate statistical techniques), presented, interpreted, stored, and secured, to use such data to solve practical problems, and for future purposes. In this regard, Ezeagu (2013), posited that data management is the effective coordination of people, equipment, and procedures together, to sort, analyze, evaluate, and distribute needed, timely and accurate data for decision-making. Data management is an administrative process that includes acquiring, validating, storing, protecting, and processing required data to ensure the accessibility, reliability, and timeliness of the data for its users (Ngdata as cited in Owan & Bassey, 2018). According to Soler, Ort, and Steckel (2016), research data are regarded as all information collected and observed for purposes of analysis and validation of original research results. Thus, data management is a deliberate and systematic process of unifying human and material resources to reach set goals. It involves regulating the activities and operation of men, materials, machines, money, and time to attain set objectives. Data management observably follows the stages of i) editing the data before being inputted ii) data inputting itself ie., inputting data from hard to soft copies iii) editing the data to be sure of consistency and credibility of the data; all of these

16870

constitute what we call managing data. If data are credible, whatever decision-making and policy arising from there, will be dependable.

In Nigeria and most of the world, higher educational institutions are serious about quality research output to promote the image of the university. This underscores the need for effective Research Data Management (RDM). Generally, RDM involves all processes and activities required to properly organize, document, store, and archive data for future use. Dimensions of Data management be seen as follows;

Figure 1: diagrammatic representation of the dimensions of data management dimensions
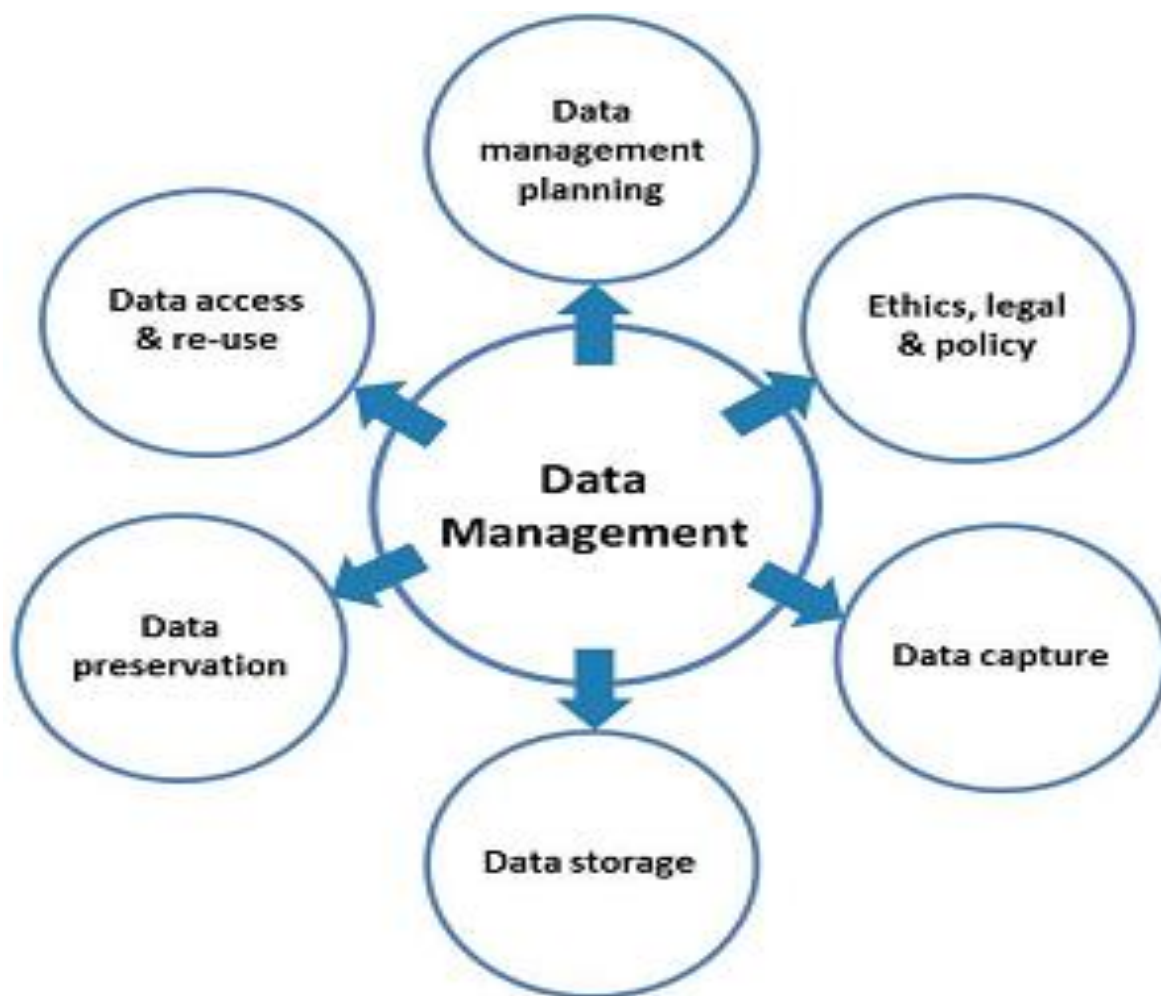
Source: Retrieved: https://dmp.qut.edu.au/faq January, 11 the 2023

In light of the above researchers must ensure the need for effective use of the dimensions of an improvement in data management. Recognizing the importance of research data and proper RDM, governments, and funding organizations encourage researchers to properly store and share data. This implies that effective data management involves the collection, presentation, and

verification of research results accomplished by researchers and scholars and thereby making it easier for other researchers to build on the existing research (Sutton, 2017).

Tripathi, Shukla, and Sonkar (2017) rightly informed that in the research world, research data are commonly seen in various areas of academic endeavor such as images, text, recordings, simulations, verbal communications, etc. They are generated when scholars (students and researchers) embark on research projects, theses, dissertations, seminars, and articles for publication. In this regard, research data involve diverse data sources depending on the research problem, the origin of data, and the discipline of the researcher. In education, observation, interview, questionnaire, checklists, inventory, and other forms of attitude and aptitude scales are employed in data collection and management. Also, in the physical sciences, experiments, observations, and computer modeling life are usually employed to collect data whereas survey questionnaires and interviews are mostly used in the social sciences. Data management planning constitutes one of the important research funders' requirements for research grant applications (Davidson 2016). However, this has made Liu and Ding (2016) note that some research universities and organizations consider data management planning as an integral part of data management practices even without any supporting policies from governments or research funding agencies. It requires flexibility and functionality that allow 'researchers to store, access, and share their data during research collaborations' (Jones, S, Pryor & Whyte, 2013). Studying existing data storage practices in the institution (such as how much data is produced; where is the data stored and what backup facilities are in place) provides a stepping stone to address problems associated with managing active data.

Smits and Teperek (2020) in their study provided an analysis of how sixteen recently graduated master's students from the Netherlands perceived research data management. It is important to study the master's students' attitudes towards this, as students in this phase prepare themselves for their careers. Some of them might become future academics or policymakers, thus, potentially, the future advocates of good data management and reproducible science. In general, students were rather unsure of what 'data interpretation' meant and would often confuse it with data analysis, study design or methodology, or ethics and privacy. When students defined the concept, they focused on privacy aspects. Furthermore, their study programmes had diverse approaches to data management education. Most of the classes offered were limited in scope. Nevertheless, the students seemed to be aware of the importance of data management and were willing to learn more about good data management practices. This report helps to catch an important first glimpse of how master's students (from different scientific backgrounds) think about research data management.

The adequate management of research data has always been an indispensable element of trustworthy scientific research, but the interest in research data management practices, skills, and experiences has flourished in the last decade. This increased recognition of data management and the sense of urgency attached to it (Feijen 2011), is partly fuelled by questions about research

16872

reproducibility and the perceived existence of a reproducibility crisis in research. Baker (2016) surveyed 1,500 researchers and the finding revealed that 90% of researchers felt that there was a reproducibility crisis (Baker 2016). It was found that, in some disciplines, over 80% of respondents experienced problems when reproducing other people's results. The survey also investigated the reasons behind such irreproducible research. In addition to a toxic publication culture ('selective reporting', 'pressure to publish' etc.), respondents also referred to (a lack of) data availability, indicating that 'methods and code were unavailable' and that 'the raw data from the original laboratory was not available'. In a recent study on Molecular Brain (Miyakawa 2020), the author noted that 'more than 97% of the 41 manuscripts did not present the raw data supporting their results when requested by an editor'. To prevent such flaws and to increase the reproducibility of scientific studies, reliable data management throughout the entire research cycle is essential.

A similar study by Owan and Bassey (2019) sought to examine data management practices that should be adopted by scholars in maintaining the quality of their research data. In achieving this, various concepts related to this paper were clarified. Various data management practices were also discussed beginning from data generation to data shredding. Based on the underlying observations from the light of the discussions made in this paper, it was recommended among others that: higher educational institutions in every part of the world should endeavour to establish a data management unit that will be saddled with the primary duty of formulating research data management policies, and the hosting of research data; every journal should as a matter of compulsion, require the submission of dataset corresponding to empirical papers submitted by authors and scholars.

Mahdi and Rafea (2019) intended to examine RDM in Iraqi universities, identify the current challenges of RDM, and proposed influential RDM practices. Data collection employed a self-administered questionnaire distributed to 155 postgraduate students and 20 faculty members from five universities in Iraq. Research findings revealed that there is a lack of proper RDM. Postgraduate students and researchers were managing their research data. The main challenges of maintaining a good RDM involved a lack of guidelines on effective RDM practices, insufficient human resources, technological obsolescence, insecure and inefficient infrastructure, lack of financial resources, absence of research data management policies, and lack of support by institutional authorities in addition to researchers being negatively influenced on research data management. Postgraduate students and researchers recommend building research data repositories and collaborating with other universities and research organizations as means of improving and sustaining research data management systems in higher institutions.

Statement of the problem

Data management in higher educational institutions in Nigeria particularly in Cross River State has raised dust among research and data scientists as to the validity and reliability of data. Researchers have proven that most data used in research undertakings have been accused of

16873

allegations of scientific misconduct. That is most researchers do not follow the normal ethical consideration in generating and managing data for research purposes. Personal interaction with most schools by the researchers shows that appropriate authorization is often lacking before the collection of data commencement by most institutions. It has also been observed by researchers that proper handling, retention, and storage of data, especially those involving humans, are not properly adhered to by most research scholars. Again, there is the problem of data distrust, inconsistencies, and complexities in available data as common among most research scholars. This is because there is a lack of guidelines on good practice, inadequate human resources, technological obsolescence, and a lack of evidence in data management. Most research scholars of different fields who claim to be data analysts and managers use fictitious data in analyzing the research outputs. This practice has affected the authenticity, credibility, usefulness, meaningfulness, and appropriateness of research outputs., and this has imparted informed policies. Generalizing research findings based on false claims can be detrimental to the overall success of research undertakings in higher educational institutions in Nigeria as in other climes. With this problem, the central focus of the study is data management in research practices in higher educational institutions: Issues in the research sphere.

Purpose of the study

The purpose of this study was to examine data management in research practices in higher educational institutions: Issues in the research sphere. Specifically, the study intends to:

1. Assess the management of data among students in terms of data collection in research practices in higher educational institutions.
2. Determine the management among students in terms of data analysis in research practices in higher educational institutions.
3. Find out data management among students in terms of data interpretation in research practices in higher educational institutions.

Research questions

The following research questions were posited to guide the study.

1. How are university students managing data in terms of data collection in research practices in higher educational institutions?
2. How are university students managing data in terms of data analysis in research practices in higher educational institutions?
3. To what extent are university students managing data in terms of data interpretation in research practices in higher educational institutions?

Hypotheses

The following null hypotheses were posited to guide the study.

1. The level of data management among students in terms of data collection in research practices is not significantly low.
2. The level of data management among students in terms of data analysis in research practices is not significantly low.
3. The level of data management among students in terms of data interpretation in research practices is not significantly low.

METHODOLOGY

This study centered on data management in research practices: Issues in the research sphere. The study was guided by three research questions and hypotheses and the research design adopted for this study was ex-post facto design with a population of 790 data analysts in tertiary. The study adopted the multistage sampling technique with a sample size of 385 (50%). The instrument used for data collection was a researchers' validated instrument titled ''Data Collection Scale for researchers (DCSR) which was validated using the inter-rater reliability method by experts in the field of Education and the coefficient of internal consistency ranged from 0.74 to 0.79. The data were analyzed using descriptive statistics (mean, standard deviation, and bar charts). The findings of the study revealed that most data analyzed were cooked (not real) irrespective of the fact that the results were quite outstanding.

Presentation of Results

This unit of the study presents the results of the hypotheses. The one-sample t-test was carried out in testing the other three null hypotheses. All decisions were taken at a .05 level of significance such that a null hypothesis was rejected if the p-value associated with the computed test statistics is less than .05 and retained if the P-value was greater or equal to .05

Descriptive Statistics of Research Variables

In using the descriptive statistics for the demographic characteristics of the respondents, simple percentages and bar chart was employed as presented in Table 1.

Table 1: Descriptive characteristics of respondents of the study

Variable	Category	n	%
Gender	Male	312	59.77
	Female	210	40.23
	Total	522	100.0
Age	Below 19 years	103	19.73
	20-25 years	138	26.44
	Above 26 years	281	53.83
	Total	522	100.0

16875

Eur. Chem. Bull. 2023, 12 (Special Issue 4), 16869-16881

From Table 1, male students dominated the study with a total of 312 (59.77% followed by female students 210 (40.23%). In terms of age, the highest number of respondents were those who are above 26 years 281 (53.83%) followed by those in the age bracket of 20-25 years 138 (26.44%) while the least are those in the age bracket of below 19 years 103 (19.73%). The results are also presented in the line graph in figure 1

Figure 1: Bar chart of the descriptive characteristics of respondents of the study
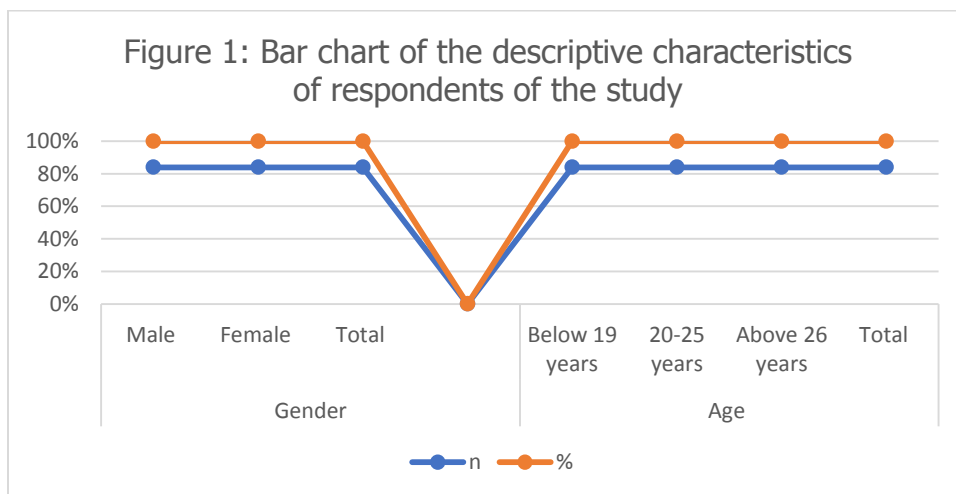
Hypothesis one

The level of data management among students in terms of data collection in research practices is not significantly low. To test this hypothesis stated. One sample t-test was applied with the level of data management (data collection) being the single variable involved. The results of the variables involved. The results of the analysis are presented in Table 2.

Table 2
The level of data management among students in terms
of data collection in research practices

Variable	n	Mean	μ	Std. Dev.	Std. Error Mean	df	LS	t-cal	p-value
Level of data management (data collection)	519	18.6859	6.00	3.98109	.17475	518	.05	106.929	.000

*p<.05

The result in Table 2 shows that the p-value (.000) associated with the computed t- value (106.929) is less than the chosen level of significance (.05). based on this outcome, the null hypothesis is rejected. This means that there the level of data management among students in

16876

terms of data collection in research practices is significantly low. As such most data collected by students are fictitious with a low level of validity and reliability.

Hypothesis two

The level of data management among students in terms of data analysis in research practices is not significantly low. One sample t-test was applied with the level of data management (data analysis) being the single variable involved. The results of the variables involved. The results of the analysis are presented in Table 3.

Table 3

The level of data management among students in terms

of data analysis in research practices

Variable	n	Mean	μ	Std. Dev.	Std. Error Mean	df	LS	t-cal	p-value
Level of data management (data analysis)	519	15.4855	6.00	3.88086	.17035	518	.05	90.9094	.000

*p<.05

The result in Table 3 shows that the p-value (.000) associated with the computed t- value (106.929) is less than the chosen level of significance (.05). based on this outcome, the null hypothesis is rejected. This means that there the level of data management among students in terms of data analysis in research practices is significantly low. Aptly put, most analyses done by students in the chosen school may look beautiful in presentation but are drawn from faulty generalizations due to the low reliability of the data collection procedure adopted

Hypothesis three

The level of data management among students in terms of data interpretation in research practices is not significantly low. One sample t-test was applied with the level of data management (data interpretation) being the single variable involved. The results of the variables involved. The results of the analysis are presented in Table 4.

16877

Eur. Chem. Bull. 2023, 12 (Special Issue 4), 16869-16881

Table 4

The level of data management among students in terms

of data interpretation in research practices

Variable	N	Mean	μ	Std. Dev.	Std. Error Mean	df	LS	t-cal	p-value
Level of data management (data interpretation)	519	18.7823	6.00	4.55562	.17035	518	.05	93.923	.000

*p<.05

The result in Table 4 shows that the p-value (.000) associated with the computed t- value (93.923) is less than the chosen level of significance (.05). based on this outcome, the null hypothesis is rejected. This means that there the level of data management among students in terms of data interpretation in research practices is significantly low. When the data collection is faulty, the analysis will also be faulty and the data interpretation will also be faulty.

DISCUSSION OF RESEARCH RESULTS

The study revealed that the level of data management among students in terms of data collection in research practices is significantly low. Thus, accurate data collecting is crucial to preserving the integrity of research, regardless of the subject of study or preferred method for defining data (quantitative, qualitative). Errors are less likely to occur when the right data-gathering tools are used (whether they are brand-new ones, updated versions of them, or already available). The present findings agree with the study of present finding agrees with the study of Owan and Bassey (2019) found that there is a need for higher educational institutions in every part of the world should endeavour to establish a data management unit that will be saddled with the primary duty of formulating research data management policies, and the hosting of research data; every Journal should as a matter of compulsion, require the submission of data set corresponding to empirical papers submitted by authors and scholars.

The results of hypothesis two showed that the level of data management among students in terms of data analysis in research practices is significantly low. Data management in terms of analysis is very important in any research experiment because it occupies a central place in making decisions based on findings resulting from the analysis of such data. Given its central

16878

Eur. Chem. Bull. 2023, 12 (Special Issue 4), 16869-16881

role, it follows that such an important asset as data deserves effective management to protect its integrity and provide an opportunity for effective problem-solving. The result aligns with the study of Mahdi and Rafea (2019) whose research findings revealed that there is a lack of proper RDM. Postgraduate students and researchers were managing their research data. The main challenges of maintaining a good RDM involve a lack of guidelines on effective RDM practices, insufficient adequate human resources, technological obsolescence, insecure and inefficient infrastructure, lack of financial resources, absence of research data management policies, and lack of support by institutional authorities and researchers negatively influenced on research data management.

The results of hypothesis three revealed that the level of data management among students in terms of data interpretation in research is significantly low. Data interpretation help in giving an adequate explanation of the generated data. The finding agrees with the study of Smits and Teperek (2020) study provides an analysis of how sixteen recently graduated master's students from the Netherlands perceive research data management. It is important to study the master's students' attitudes towards this, as students in this phase prepare themselves for their careers. Some of them might become future academics or policymakers, thus, potentially, the future advocates of good data management and reproducible science. In general, students were rather unsure of what 'data interpretation' meant and would often confuse it with data analysis, study design or methodology, or ethics and privacy. When students defined the concept, they focused on privacy aspects.

CONCLUSION

In the research environment, data management occupies a central role for everyone who needs academic growth. conducting quality research demands quality data collection which metamorphoses into quality data analysis and interpretation of the outstanding finding. When data management is ignored in any research endeavor there is the likelihood of false generalization. The study, therefore, concludes that the level of data management among students in terms of data collection in research practices is very low. Also, the level of data management among students in terms of data analysis in research practices was found to be very low. Though quite a few research exercise done by students tends to be accurate, they still retain a low level of reliability. Finally, the level of data management among students in terms of data interpretation in research practices is very low among the class of respondents sampled. Against these odds, the study recommends that institutions of higher learning should focus more on data management in research practices.

RECOMMENDATIONS

Based on the findings of the research study, the following recommendations were posited.

1. School authorities at all higher education institutions should have a data collection bank to help screen research data which will help to improve research data management practices at

16879

national levels. This is because most data were analyzed by quacks who do not have the prerequisite skills in research practices.

2. In the higher educational institutions in Nigeria, research development centers should be mounted along with research ethics committees and other data protection organs to help plan, manage, and safeguard research data. Research from various units can be submitted to the Research Ethics Committee to determine whether it has been compromised or not. Other tools for discovering compromised data – Open data kit, Plagiarism checks, Generation-evaluation follow-up -re-evaluation – RDM System -Open access, etc., can be deployed and made popular in their use to safeguard research data. The study recommended the need for compulsory participation in seminars and conference participation for all lecturers and students in the academic environment.

3. RDM plans should be made mandatory as a prerequisite for NUC accreditations of programmes which need to be mandatory by universities for its grantees with follow-up on maintaining updates of its specifications annually. It was, however, concluded that the need for data management centers to be established by schools. This will help checkmate and promote quality data management by research scholars.

REFERENCES

Baker, M. 2016. 1,500 Scientists Lift the Lid on Reproducibility, 25 May 2016. Nature News. [Last accessed 30 April 2020]. DOI: https://doi.org/10.1038/533452a.

Davidson, J 2016 Introduction to data management planning. , p.23. Available at: http://www.dcc.ac.uk/resources/publications/2016 [Accessed January 13, 2017].

Feijen, M. 2011. What Researchers Want. SURF Foundation. Available at: https://www.bvekennis.nl/wp-content/uploads/documents/11-0198-What-researchers-want.pdf [Last accessed 30 April 2020].

Idika, D.O., Joshua, M.T. & Umoinyang I.E (2017). Assessment of research skills among academic staff in Public universities in Akwa Ibom and Cross River State, Nigeria. Nigerian journal of educational research and evaluation. 16 (2), 2-18.

Jones, S, Pryor, G & Whyte, A (2013). How to Develop Research Data Management Services - a guide for HEIs. Digital Curation Centre, (March), pp.1–22. Available at: http://www.dcc.ac.uk/resources/how-guides.

Liu, X. & Ding, N. (2016). Research data management in universities of Central China: practices at Wuhan University Library Introduction. The Electronic Library, 34(5). Available at: http://www.emeraldinsight.com/doi/pdfplus/10.1108/EL-04-2015-0063 [Accessed May 14, 2017].

16880

Eur. Chem. Bull. 2023, 12 (Special Issue 4), 16869-16881

Mahdi, S.M & Rafea, I. (2019). RDM in Iraqi Universities, identify the current challenges of RDM, and propose influential RDM practices. DESIDOC Journal of Library & Information Technology, 39, (6), 308-314.

Miyakawa, T. 2020. No Raw Data, No Science: Another Possible Source of the Reproducibility Crisis. Molecular Brain, 13(24). DOI: https://doi.org/10.1186/s13041-020-0552-2.

Owan, V. J. & Bassey, B. A. (2019). Data management practices in educational research. In P. N. Ololube & G. U. Nwiyi (Eds), Encyclopedia of institutional leadership, policy, and management: A handbook of research in honour of Professor Ozo-Mekuri Ndimele (pp. 1251 – 1265). Port Harcourt, NG: Pearl Publishers International Ltd.

Smits, DAB and Teperek, M. (2020). Research Data Management for Master's Students: From Awareness to Action. Research Data Management for Master's Students: From Awareness to Action. Data Science Journal, 19 (30), 1–11.

Sutton, C. (2017). Managing and sharing research data: A guide to good practice. London Sage ISBN 978-1-446- 26726-4". J. Royal Stat. Soc.: Series A (Statistics In Society), 180 (3), 940-941.

Tripathi, M.; Shukla, A. & Sonkar, S. K. (2017). Research data management practices in university libraries: A study. DESIDOC J. Libr. Inf. Technol., 37 (6), 417-152.

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

16881

Eur. Chem. Bull. 2023, 12 (Special Issue 4), 16869-16881