Research Data Management Practices of Chemistry Researchers in Federal Universities of Technology in Nigeria

Abstract

This is an exploratory study on the research data management practice of chemistry researchers in federal universities of technology in Nigeria. Research data are increasingly becoming important and are required by journal publishers and international funders to be carefully gathered and properly kept for validation of results, sharing and for future reuse. These chemistry researchers were purposively selected across the five federal universities of technology in Nigeria namely: Federal University of Technology, Minna (FUTMIN); Federal University of Technology, Akure (FUTA); Federal University of Technology, Owerri (FUTO); Abubakar Tafawa Balewa University of Technology Bauchi (ATBU); and Modibbo Adama University of Technology, Yola (MAUTECH). A qualitative research technique using semi-structured interview schedule was used to gather data from forty (40) chemistry researchers across the five federal universities of technology in Nigeria. These chemistry researchers were those with either a doctoral degree or currently undergoing a doctoral training across all available options and ranks in chemistry. Interview questions were divided into four major sections covering chemistry researchers' understanding of research data, experience with data loss, data storage method and backup techniques, data protection, data preservation, and availability of a data management plan. Findings, as derived from thematic analysis of transcribed responses, revealed that chemistry researchers' research data management practices needs a lot of improvement to ensure long-term access, preservation and reuse of their research data. The study recommended the need for libraries to embark on awareness and education of chemistry researchers on the benefits of good research data management practices.

Keywords: research data, research data management, chemistry researchers, Federal University of Technology Minna, Federal University of Technology Akure, Federal University of Technology Owerri, Modibbo Adama University of Technology Yola, Abubakar Tafawa Balewa University Bauchi, Nigeria.

Introduction

Appropriate research data management practice ensures that underlying research data necessary for validating and supporting research findings are available and usable by the scientific community. The practice also prevents irrecoverable loss of data and information in the long run. Research data, as an integral part of a research, are being required by impact-factor journals to be submitted alongside articles before they are accepted for publication. This is to enable research findings to be verifiable and reused. Government and international funders also see research data as a return-on-investment because another user can re-purpose and reuse it. Giving access to research data for reuse by secondary users ensures that new researches are novel and not re-inventing the wheel or duplicating previous researches.

Research Data Management is a standard part of good research practice offering a lot of benefits to the society, research funders and the research community. Jones *et al.* (2013) gave the Digital Curation Centre (DCC) definition of Research Data Management (RDM) as "the active management and appraisal of data over the lifecycle of scholarly and scientific interest". It can help researchers by potentially increasing efficiency, saving time and resources and boosting the impact and visibility of their work through openness and transparency.

A study by Smith (2014) observed that the challenge of how to effectively manage research data affects all scientists and researchers within and across multiple domains. The study observed further that most researchers struggle unsuccessfully with storage and management of their burgeoning volume of documents and datasets that they need and that result from their work. The consequent of this is that rising accumulation of useful findings may be lost or unavailable when conducting future research. There is therefore the need to incorporate a good data management practice throughout the research workflow.

Chemistry is the study of matter and energy and the interaction between them. It is sometimes referred to as 'central science' because it connects other sciences to each other, such as biology, physics, geology and environmental science (Helmenstine, 2016). This scientific field of study is of different options that researchers may wish to specialize. These options are analytical, organic, inorganic, polymer, nano, environmental, industrial, physical, petroleum, computational and food chemistry. Each of these options involve various stages of experiments that generate distinct research data. Managing these research data becomes necessary and challenging especially now that it is being required to be submitted alongside articles for publication in impact factor journals and it is also a requirement for applying for international research grants and funding.

The aim of this study was to explore the research data management practices of chemistry researchers throughout their research process. The study provided chemistry researchers' insights in the following areas: understanding of the concept of research data, research data loss experiences, preferred research data storage media, research data protection method, research data preservation, and awareness of data management plan.

Literature Review

Libraries and Research Data Management Services

Academic libraries around the world are rolling out research data management services to assist and guide researchers in safeguarding data generated during research activities and to also ensure that such data are reusable and preserved for future reference. This, according to Cox & Pinefield, 2014; Koltay, 2016) as cited in Tenopir et al. (2017), is because academic libraries by their nature are critical stakeholders in research data preservation and management now and in the future. Also, academic libraries can also roll out research data services as part of efforts aimed at enhancing research workflow in African universities (Abduldayan, et al., 2016). Cox et al. (2017) studied the developments in research data management in academic libraries using the "landscape maturity model" that reflected the current and planned research data services and practices in academic libraries in developed countries. A survey designed in the form of a questionnaire was used and administered online to a large number of academic libraries in the UK, Australia, Canada, Germany, Ireland, Netherlands, and New Zealand. Findings revealed that libraries are increasingly providing advisory and non-technical support services as well as policy development around research data management (RDM). However, the study noted the limited support for technical RDM services such as provision of data catalogue and curation of active data.

Libraries are tasked with the duty of preserving research data because losing data jeopardises the future of research scholarship, hence, librarians should be the custodian of not just intellectual contents but also datasets. In a follow-up survey of institutional support services for RDM, Tenopir et al. (2017) also presented the results of a survey of directors of the Association of European Research Libraries (LIBER) academic member libraries to discuss what types of research data services are being offered and what services are planned for the future. The study revealed that efforts are in place for staff RDM skills development. The findings also revealed the need for further study on the different research data management needs of researchers from various disciplines. European libraries are reported to be collaborating with many internal and external partners as well as working on developing policies towards effective research data services.

Librarians and Research Data Management Services

The role of librarians are shifting from traditional information resources manager to actual data manager, otherwise known and referred to as data curators. Tammaro et al. (2018) in their study of "data curators roles and responsibilities: an international perspective" noted that this new role require librarians and information professionals to have the requisite knowledge of the research process and also skills in managing and curating data generated throughout the research lifecycle. The study emphasized further on the importance of research background for data curators, librarians and information professionals and its implications for education of future data experts that may need to combine technical skills, expertise in metadata and information organisation standards, and knowledge of research process and methods. According to the study, a participant noted that it is easier for data curators to build credulity and establish trust with researchers when their existing research lifecycle is understood. This helps to ensure that the RDM discussion is centred on what the researchers are already familiar with.

Research Process of Chemistry Researchers

Research data are generated at every stage of any scientific research process. Hence, understanding the chemistry researchers' research process is important so that their research data management practices can be properly defined, understood and presented. Since RDM is a new term especially to chemistry researchers in Nigeria, and an emerging aspect of research best practices, the generic research process (shown in Figure 1) helped to guide the RDM discussions with chemistry researchers as the authors initially tried to understand the existing research process in chemistry before starting the RDM conversations with the chemistry researchers. This is in line with Tammaro et al. (2018) study that encouraged understanding research processes and methods so as to be able to 'talk the research talk'.

Research process in chemistry is often started with a problem statement, followed by an objective and finally, an expected outcome. These expected outcomes are sometimes stated as a hypotheses to be tested with an experiment. Research methodology differs across options although there are areas of overlap among them. Software is used for research design and interpretation of result and concluded research are usually disseminated through publications, conferences, commercialised or patented. Research data are generated right from the beginning of the process till the end when the study is finally concluded.

Figure 1 shows the generic research process of chemistry researchers in Nigeria:

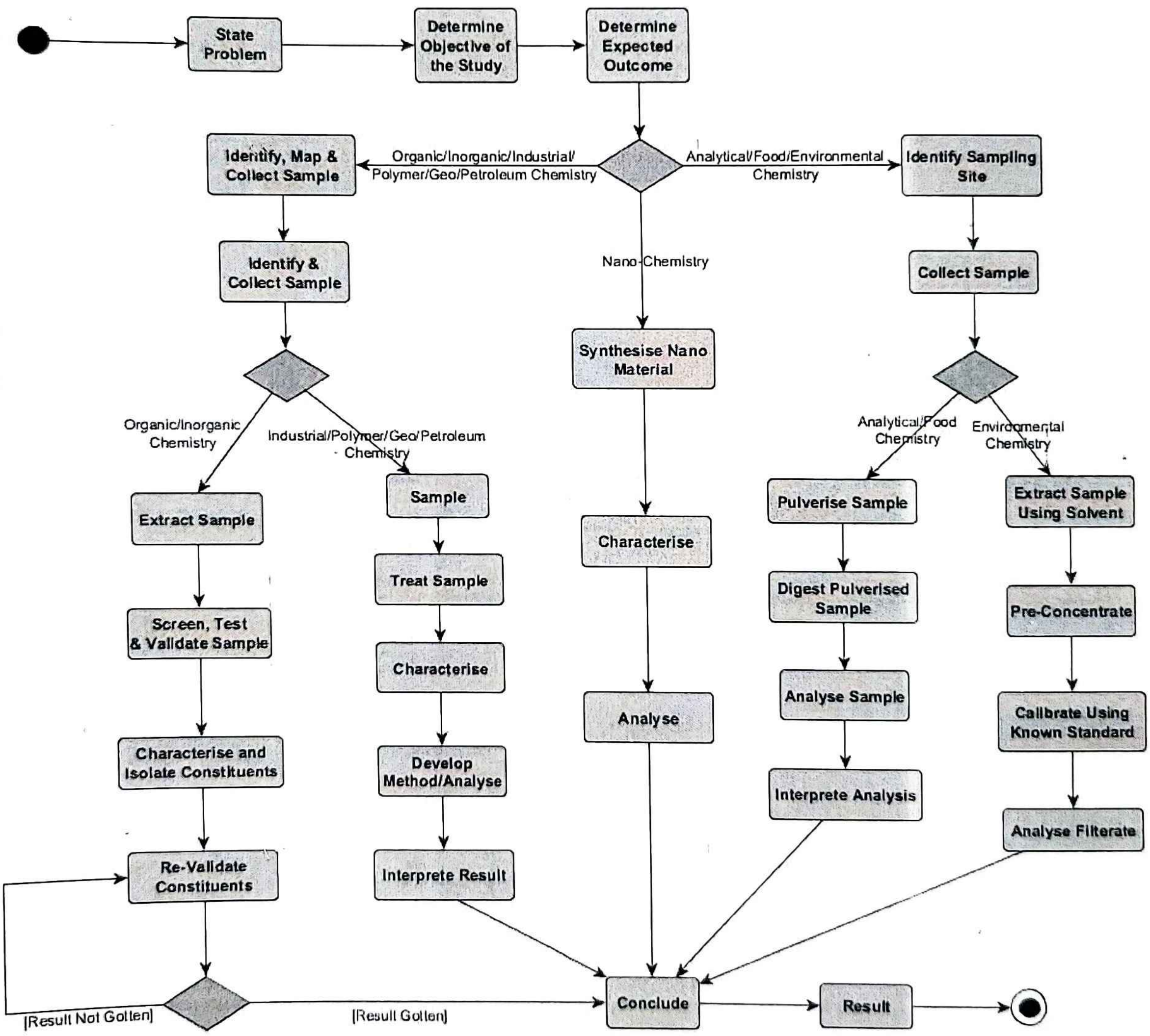

Figure 1: The Generic Research Process of Chemistry Researchers in Nigeria (Author's Concept)

Several studies (Van Gulick & Borgi, 2018), Aydinoglu et al. (2017), Tripathi et al. (2017), and Tenopir et al. (2011) have explored data management practices and perceptions of researchers in their respective institutions. According to Tripathi et al. (2017), a brief assessment of researchers' perceptions towards research data management in India revealed that the library have a major role to play in championing the effective management of raw research data. This study interviewed forty (40) researchers and faculty members from two universities in India with a view to understanding how they handle raw data. The study recommended the need for a national level policy on RDM and a reward system for researchers that comply with the initiative.

In Turkey on the other hand, Aydinoglu et al. (2017) there is a lack of research data policy or strategy, hence RDM does not exist from an institutional perspective. The major Turkish funding agency does not have a defined data management requirement that researchers must abide with neither does the institutions have an RDM policy. The study also revealed that Turkish researchers lack the required skills and knowledge needed for RDM however they are in support of data sharing with colleagues under certain conditions.

Chen & Wu (2017) conducted a survey on the needs for chemistry research data management and sharing among 119 researchers and students of chemistry in the Chinese Academy of Science. The study provided an insight into the attitudes and needs of chemistry students on data management and sharing to enable the library develop a range of RDM services for the chemistry research community in China.

Only few studies (Chiware & Mathe, 2015) and (Raju & Schombee, 2013) are on data management practices and perceptions of researchers in Africa. There is none in Nigeria as at the time of this study, a gap this study hopes to fill. Although some findings of this study show areas of similarities and differences in data management practice of Nigerian researchers and their counterparts abroad, this study presents the actual RDM practices of chemistry researchers in federal universities of technology in Nigeria.

Methodology

The research design used for this study was the exploratory research design involving qualitative research instrument. A semi-structured oral face-to-face interview was conducted with selected chemistry researchers across the five federal universities of technology in Nigeria. Interview method was adopted to allow chemistry researchers express themselves clearly on their research data management practices. These federal universities of technology are: Abubakar Tafawa Balewa University of Technology, Bauchi (ATBU); Federal University of Technology, Akure (FUTA); Federal University of Technology, Minna (FUTMIN); Federal University of Technology, Owerri (FUTO);. and Modibbo Adama University of Technology, Yola (MAUTECH) spread across three geopolitical zones in Nigeria. Interview method was adopted to allow chemistry researchers express themselves clearly on their practices of research data management.

Interviews lasted for approximately thirty minutes while the least response time was about nine minutes depending on the respondents' knowledge of the questions asked. Research Assistant was employed to assist with the audio recording and to ensure that informed consent forms were properly filled and collected. All interviews were conducted in the respondents' workplace after an agreed day and time which is usually at about 9am-4pm of any working day. The data gathering process took approximately ten weeks (2months and 14days) due to the distance involved in traveling across the four geopolitical zones in Nigeria where the universities were located.

Total enumeration technique was used to cover all chemistry researchers in the departments of chemistry in all federal universities of technology in Nigeria. Then, stratified sampling method was used to divide the population into different subgroups (strata) and to select subjects from each stratum in a proportionate manner (Dudovskiy, 2015). Chemistry as a field of study has various options or areas that researchers can specialize on. These options include: industrial chemistry, organic chemistry, inorganic chemistry, environmental chemistry, analytical chemistry, physical chemistry, polymer chemistry, and nano chemistry. Forty (40) chemistry researchers were interviewed orally in their workplace. These 40 chemistry researchers covered all available options in chemistry.

The interview schedule used for understanding data management practice of chemistry researchers was divided into four sections and questions were asked to provide insights in the following areas: understanding research data, research data loss, research data storage media, research data protection, research data preservation, and data management plan. For the data transcription and analysis, this study adapted the Braun & Clarke (2006) thematic analysis approach which involves a six-phase of qualitative data analysis. These phases are: transcribing data, generate initial code, search for themes, review themes, define and name themes, and finally present the final report of the analysis. In addition, the *Provalis* Qualitative Data Analysis (QDA) Miner (version 5) software was used for generating themes and subthemes from the coding framework and the final report was presented in Figure 2.

Presentation of Results

Table 1 shows the breakdown of chemistry researchers that participated in the interview session:

Table 1: Breakdown of Respondents into Rank and Options Strata

S/N	University	Rank of Respondents	Options in Chemistry	Number of Respondents
1.	ATBU	Professor – 1 AP/Reader – 3 Senior Lecturer – 2 Lecturer I – 1	Organic, Environmental, Physical, Inorganic, Analytical.	7
2.	FUTA	Professor – 1 AP/Reader – 1 Senior Lecturer – 2 Lecturer II – 3 Assistant Lecturer – 1	Food, Environmental, Petroleum, Geochemistry, Analytical, Industrial, Polymer Chemistry.	8
3.	FUTO	Professor – 2 AP/Reader – 1 Senior Lecturer – 3 Lecturer II – 2	Organic, Analytical, Physical, Environmental, Polymer, Computational Chemistry.	8
4.	FUTMIN	Professor – 1 AP/Reader – 3 Senior Lecturer – 3 Lecturer I – 2 Lecturer II – 2	Analytical, Polymer, Organic, Environmental, Nano chemistry.	11
5.	MAUTECH	Professor – 1 AP/Reader – 2 Lecturer I – 2 Lecturer II – 1	Polymer, Physical, Organic, Environmental, Analytical.	6
	Total			**40**

Figure 2 shows the major themes derived from the data analysis. These themes were further discussed under the following sections:

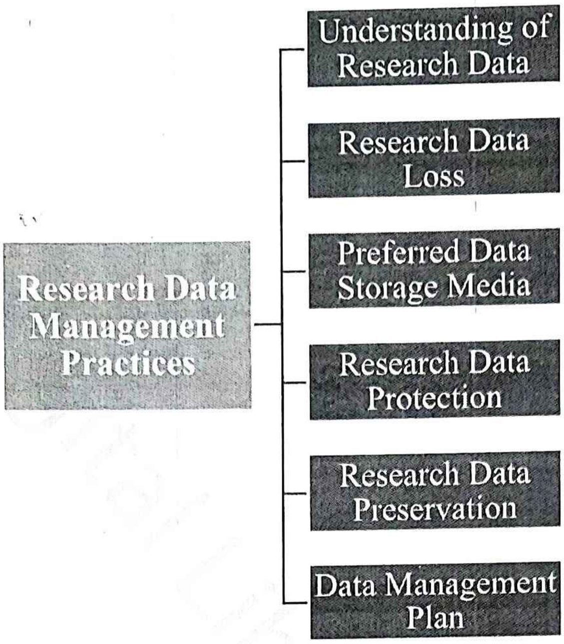

Figure 2: General thematic diagram of Data Management Practices of Chemistry Researchers in FUTs in Nigeria

The interview guide can be found in Appendix A while Appendix B presents an overview of responses by chemistry researchers' on their research data management practices. The differences and similarities in opinions of chemistry researchers across the five federal universities in Nigeria was also shown on the table in Appendix B.

1. Understanding of Research Data

Chemistry researchers in FUTs in Nigeria define data and research data based on their understanding of the concepts. In ATBU Bauchi, chemistry researchers see data as a sequence of numbers or figures derived from experiments which can be used for validation. Research data are however seen as data from research work. Chemistry researchers in FUTA see data as collection of facts gathered from experiments while research data are generated from such experiments conducted in a laboratory. Chemistry researchers from FUTMIN defined data as results gathered from investigations while research data is seen as facts acquired from research. Chemistry researchers in FUTO defined data as results of experimental studies presented in figures, results of experiments, collection of information gathered to enhance research, facts, quantifiable results as well as results obtained from quantitative or qualitative analysis. Research data however is seen as results obtained from research studies in the laboratories, or from a systematic study which could be intermediate or final derived during the research. Chemistry researchers in MAUTECH defined data as valuable resources to get results, a set of figures, information collected for the purpose of studying, any information in bits, groups, or quantity that can be processed further. Research data on the other hand is seen as outcome from research, set of figures collected in the field, data obtained from experiment that is aimed at getting results, and data used for finding out things.

A chemistry researcher from FUTA sees research data as:

> "research data? I know you cannot get data except you make a research"

However, chemistry researchers generally agreed that research data is:

> "any information gathered before, during and after an experiment or research",
>
> "... outcome of experiments or investigations emanating from research for further studies",

and that it can be in different forms- pictures, numbers, figures, texts, graphs. In ATBU Bauchi, chemistry researchers' research data are in the form of experimental data, field data, and observation, physical and chemical data. FUTA chemistry researchers' research data types are laboratory data, environmental data, and primary data. In FUTMIN, research data could be experimental data, primary, secondary, plant, human, geographical, field, or statistical data. FUTO research data types are laboratory data, experimental data, conceptual data, simulation data, field, empirical, and data from questionnaire. The different types of research data according to MAUTECH chemistry researchers include plant data, field data, laboratory data, experimental data, questionnaire, interview, and archived data. This shows that a large percentage of chemistry researchers have an understanding of what research data is and the most common research data type is experimental data followed by laboratory data, questionnaire data, interview data, field data, and plant data.

Knowledge of Digital Data

ATBU chemistry researchers generate and use digital data based on the research being carried out. These digital data are mostly in the form of graphs, images, MATLAB, and NMR Spectra. Chemistry researchers in FUTA also generate and use digital data. Their digital data are in the form of spreadsheets, texts, images, NMR spectroscopy, and graphs. FUTMIN chemistry researchers generate digital data which are usually in the form of spreadsheet, text, image, video, Mat lab, NMR spectroscopy, graphs, Thin Layer Chromatography (TLC), and other customized files from instrument-specific software. For FUTO chemistry researchers, digital data generated could be in the form of graphs, spreadsheets, images, tables, MATLAB (matrices), and texts. Chemistry researchers in MAUTECH all generate digital data which were mostly in the form of spreadsheets, images, figures, charts, NMR spectra, scanned documents, and graphs.

All chemistry researchers are knowledgeable about digital data which are usually in the form presented in the word cloud below. Spreadsheets and NMR spectra are the most common types of digital data generated by chemistry researchers in FUTs in Nigeria. This is shown in Figure 3.

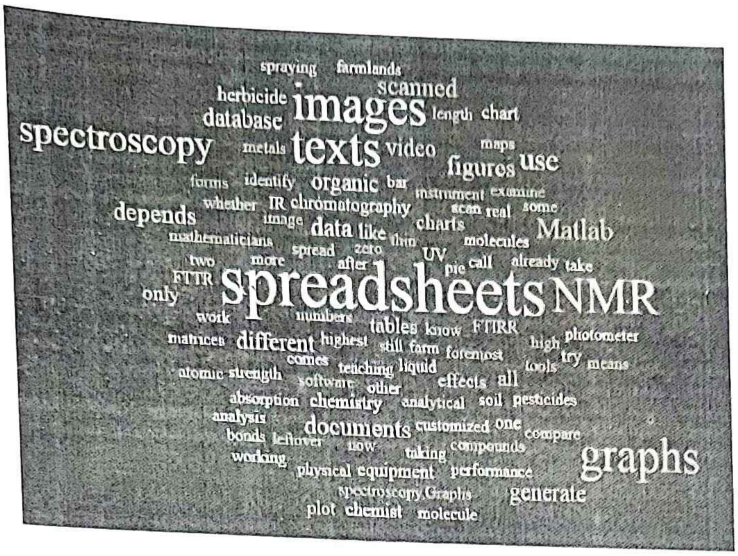

Figure 3: Word Cloud showing Types of Digital Data

Software Used for Digital Data Generation

In ATBU Bauchi, some of the software used for digital data analysis are Minitab, SPSS, Origin 8, Origin 10, CHEMSK, Material Studio, ACCRYLS. In FUTA, software mostly used for analyses are Microsoft Excel, SPSS, and NMR spectrometry. In FUTMIN, software used for digital data generation and processing include Origin, Minitab, Matlab, SPSS, MS Excel, ChemDraw, PLS, QSAR and MS Word. Chemistry researchers in FUTO uses Origin, Microsoft Excel, OriginPro, MATLAB, electrochemical data equipment software, OGR, SPSS, ArchView, AquaChem, MudFlow, QTIPLOT, Zigmaplod, and Artificial Neural Network (ANN) software. In MAUTECH, software used for data generation are ChemDraw, ChemSketch, UC Spectra, instrument-specific software, SPSS, UC machine software, Aurora.

SPSS is the most common software used for digital data generation as depicted in the word cloud in Figure 4.

Figure 4: **Word Cloud showing Software for Digital Data Generation and Analysis**

Data Naming System

Chemistry researchers in ATBU Bauchi store and name data using the source, name, title, codes, and date. In FUTA, data are stored and named using self-devised names, research-dependent names or just sample name. They also use defined abbreviations and codes for saving files and data. FUTMIN chemistry researchers store and name digital data files using date and day of data generation, name of researcher, name of sample, site name, codes, and any other convenient self-devised means. Chemistry researchers in FUTO do not have any specific naming convention as data are stored and named based on the current research, self-devised, use of date, time and year, researcher's middle name, student's name generating the research, instrument used for analysis, experiment-based, and location of research sample. However, to identify different versions of stored files, chemistry researchers in FUTO make use of number, day of research, storing on different CDs/flash, number of runs, dates, and current research result. For MAUTECH chemistry researchers, digital data files are stored using random self-devised methods based on the research. Dates, initials, numbers, use of children's names, acronyms, codes and condition of experiment are used to store different versions of digital data on the system. Date is the most commonly used method of saving data files. A researcher from ATBU opined that:

> *"It depends on the data I have, I just name from the source. I can use it"*

While from FUTMIN:

> *"I use data, day, name of researcher, coding"*

> *"I use my name, if I open my computer you'll see [researcher's name]2010"*

Another researcher from FUTA said:

> *"I can use anything, sometimes I can use my name, I can use the name of my son, it depends on what I want to name"*

In storing research data, devising a distinct naming style is of utmost importance to aid retrieval of data files. This sub-node uses bar chart to show methods used by chemistry researchers in saving their digital data files as shown in Figure 5.

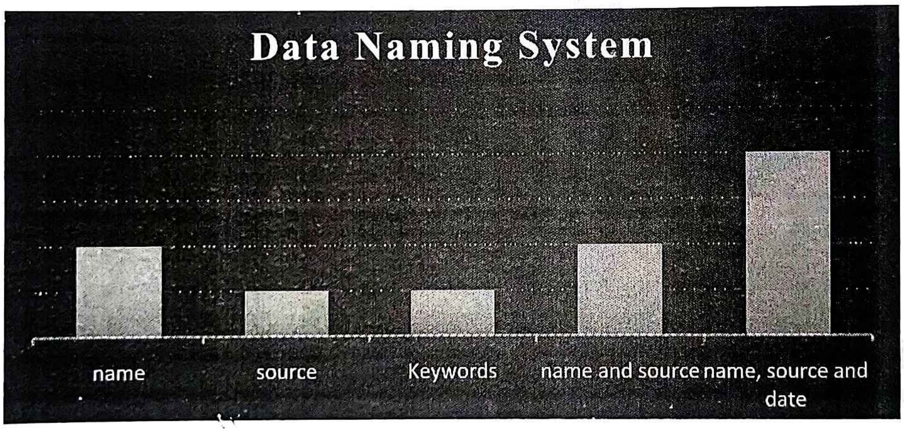

Figure 5: Data Naming System of Chemistry researchers in FUTs in Nigeria

Chemistry researchers in ATBU Bauchi manage different versions of stored data by using NMR-specific numbers, UV-generated numbers, self-devised numbers and codes. Other details of stored data files are generated by the computer automatically. To identify different versions of stored files, chemistry researchers in FUTO make use of number, day of research, storing on different CDs/flash, number of runs, dates, and current research results. Chemistry researchers in FUTMIN and MAUTECH manage different versions of data using folders and other details such as dates are generated by the computer system automatically.

2. Research Data Loss

Interestingly, chemistry researchers in ATBU agreed that they have at one time or the other experienced a data loss situation. This was as a result of storage device theft, computer system malfunction, and sudden file corruption. Although some of the chemistry researchers were able to manage the situation because they had backup on an external drive while majority said they lost such files completely. Most chemistry researchers in FUTA have had painful experiences of data loss which was as a result of not saving while working on the system, corrupted files copied from an external device, and sudden system crash. Regular backup and continuous saving while working are lessons learnt from the sad data loss experience. Chemistry researchers in FUTMIN agreed to have lost critical and often irrecoverable data due to poor storage device, system crash, bad memory card, virus attack, and system theft. Lessons learnt from data loss experience included using clear labels and names for stored data files, backup data in multiple locations, proper handling of external storage devices, seeking help from computer specialist when confronted with system faults, adopting online storage and cloud services. Only one chemist in FUTO has never experienced a data loss situation due to poor file storage, unrecognized file format or bad storage device. Others have had the unfortunate experience due to damaged external hard disc, power failure while conducting an experiment, spoilt flash drive, corrupted files, system crash, system theft, and data loss in the process of upgrading from one software or system to the other. Lessons learnt from these experiences are regular backups, constant file saving while working, and creating backup in multiple locations. All MAUTECH chemistry researchers have experienced data loss due to poor file storage, unrecognized file format or bad storage device.

"This usually happens as a result of system crash, system theft, and sudden corruption of system, improper file naming, virus attack, and system malfunction".

A chemistry researcher from ATBU shared that:

"Serious! what happen actually was as I took my data and have been recording in the storage machine then I connected the computer to another machine that is supposed to display the work from scrap and somehow I don't know what happened but it affected the computer and I could not see the data but fortunately I had some back up but not all. I was forced to go back and repeat the experiment for latter part of the work".

Another research data loss experience from a chemistry researcher in FUTA was that:

"...in fact something happened yesterday, I don't know (giggles), my existing file, all of a sudden, I just saw something like that, like Arabic something or no no Chinese, I don't know the type of language something, ahh, I hope this is not virus, I closed it and opened it and everything was gone."

Lessons learnt from the unfortunate incidence include the need for external backup, sharing, updating and storing in different places, patience while saving files, and proper security of stored data.

3. Preferred Data Storage Media

According to ATBU chemistry researchers, storage media to be used for managing research data depends on the value of work done, however, the most important factor in determining storage media is security. For FUTA chemistry researchers, the choice of storage media to use for storing data is dependent on the portability, accessibility and safety from virus of the storage media device. Factors that affect FUTMIN chemistry researchers' choice of storage device include portability, durability, accessibility, security, mobility and reliability. Preferred choice of storage device are external hard discs and online storage such as drop box, email, Google drive, depending on the volume of the work. Chemistry researchers in FUTO consider the following when deciding on the storage media to use- durability, storage capacity, and portability, susceptibility to virus, ease of retrieval, mobility, and safety. Researchers here keep record of stored files and data in office lockers, at home, with a research assistant, and in the laboratory. In deciding storage media to use for data files, chemistry researchers in MAUTECH consider cost-effectiveness, nature of the data, security of the media, ease of access, and relevance to research. Other factors include capacity of the storage media, mobility, and susceptibility to virus attack, accessibility, and compatibility. Most preferred storage devices among chemistry researchers are computer hard drive, Google Drive, flash disk, CDs, external hard disk.

Figure 6: Determinants of Storage Media Type

Knowledge on Types of Storage Media

ATBU chemistry researchers are aware of the advantages and disadvantages of some storage devices especially external hard disks, flash disks, cloud storage using personal email, and compact discs. They are however not aware of same for other cloud storage devices and services. FUTA chemistry researchers are aware of the advantages and disadvantages of various storage media. Most of FUTMIN chemistry researchers know the advantages and disadvantages of the different storage devices and media. Majority of chemistry researchers in FUTO are aware of the advantages and disadvantages of the different storage media. Only few chemistry researchers had little or no knowledge of the advantages and disadvantages of different storage media.

Estimation of Research Data Size

Chemistry researchers in ATBU revealed that their research data can grow from megabytes to gigabytes depending on the research and funding capacity. FUTA chemistry researchers opined that their data files can also grow to as large as gigabytes depending on the type of research. They also noted that estimating the growth of data files from the beginning of the research is not easy except the research is concluded. The estimated size of research data by FUTMIN chemistry researchers were mostly in megabyte, although some research data could run into gigabyte and even terabyte depending on the type of research. Most researchers in FUTO estimated that their research data can grow into gigabytes and even terabytes depending on the type of research. Only one researcher estimated research data to be in megabyte. In MAUTECH, chemistry researchers also noted that their estimated file size while conducting research might be in megabyte, although the estimate was not easy to give as it depends on the research being conducted.

Chemistry researchers were able to estimate the size of their research data using megabytes, gigabytes and terabytes. However, a large percentage of them could not predict the increase. The proportion of chemistry researchers that could predict data increase is as described in Figure 7.

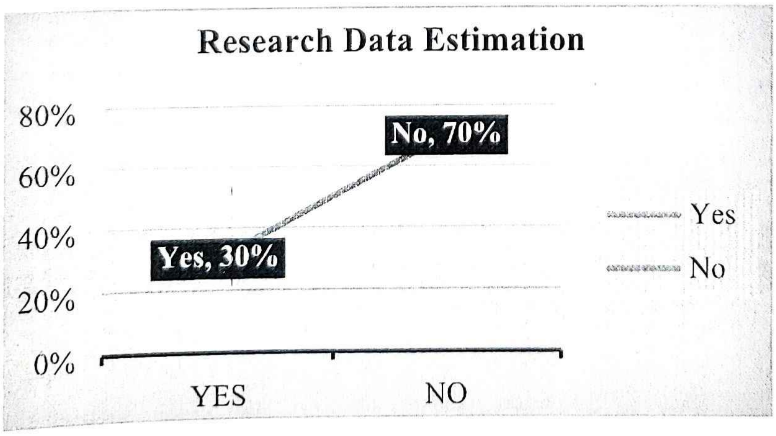

Figure 7: Estimation of Research Data Size

Backup Methods and Additional Storage Device

Backup method of ATBU chemistry researchers is to update data saved on external devices regularly. FUTA chemistry researchers perform updating and backup of data files occasionally either on the master file only or on all other files stored in various locations. In FUTMIN, chemistry researchers adopt periodic update of files on their local system, upload to cloud storage devices, use of external hard disks as well as automatic backup. Only few researchers do not engage in periodic update of files and data. Additional storage device mostly used include system hard disk, CDs/DVDs, flash sticks, hard copies and cloud storage. FUTO chemistry researchers backup their data as the research is ongoing, they also use dual saving mode, folders, and perform regular updates. External hard discs, cloud storage, flash sticks, CDs, email dedicated for research are the mostly used additional storage media while CDs is the most preferred. Flash sticks, CDs, external hard drives, and hard copies are used as additional storage media by MAUTECH chemistry researchers. Also, data are stored in different locations and on external hard drives as a backup method.

Other backup methods of chemistry researchers are displayed on the word cloud in Figure 8.

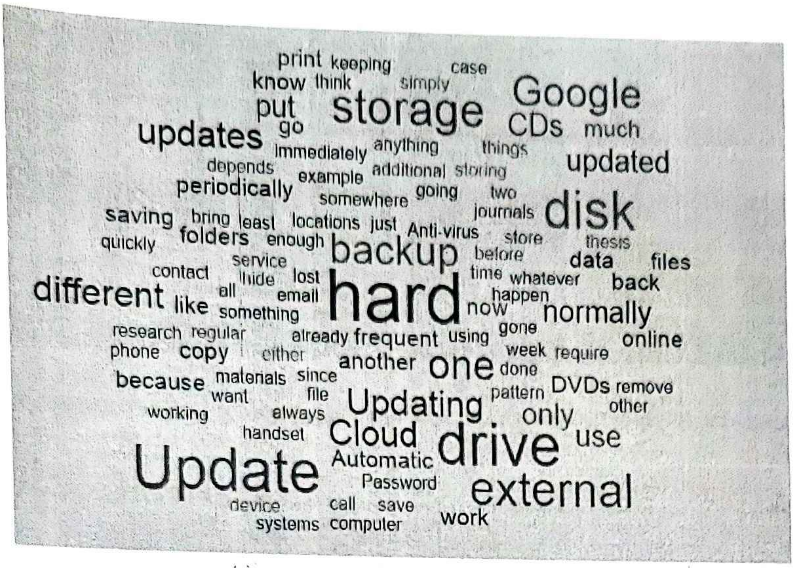

Figure 8: Backup and Additional Storage Device of Chemistry researchers

4. Research Data Protection

In ATBU, chemistry researchers secure their stored data using passwords, and use of lock and key on file cabinets in the office. In FUTA, although a researcher did not believe in using password or any form of security to protect data files on the computer system, other researchers agreed to the use of password and physical security using lock and key. FUTMIN chemistry researchers use passwords, PINs, coding, anti-virus, and periodic scanning of files as security measures to protect stored data and files. FUTO chemistry researchers keep record of stored files and data in office lockers, at home, with a research assistant, or in the laboratory. MAUTECH chemistry researchers also use password for stored digital files, while lock and keys are used for physical data files as a means of securing them from unauthorised access.

Hardware or Software Failure

Hardware and software failures are a common experience for all chemistry researchers. For ATBU chemistry researchers, such failure is usually in the form of inability to access a previously saved file after a long time or a complete system crash. Files were sometimes recovered from other researchers that they shared such files with before the crash or from the cloud storage. In FUTA, their experience is not different from the experiences shared by chemistry researchers in other FUTs in Nigeria, however, FUTA chemistry researchers usually contact the university's Information Technology Service (ITS) unit for help with data recovery from their computer systems. Software compatibility issues are experienced by most chemistry researchers in FUTMIN which usually leads to loss of data. The way out as shared by the chemistry researchers was to contact colleagues for help, update software, perform system scan, or visit computer experts for data recovery. Some chemistry researchers revealed that they simply let go of such files that are unreadable or try to get from another source. Majority of chemistry researchers in FUTO has experienced hardware or software failures and inability to access old files. They described the situation as frustrating and confusing.

> *"There was a time I was using materials studio it just did not want to come up so I had to format and lost everything"*

To regain loss data, chemistry researchers had to either consult a computer specialist, update the software, repeat experiment, perform system restore, copy from external storage device, or let go of the loss data. All MAUTECH chemistry researchers have experienced a hardware or software failure and inability to access old files which led to complete loss of data, and suffering in trying to get another data. This, they say, has taught them the importance of proper backup to avert future data loss. They however requested for advice and guidance on proper data storage and back up.

5. Research Data Preservation

Chemistry researchers were further interviewed on where they would like their data to be preserved and where to go seek for help on research data long-term preservation issues. ATBU chemistry researchers believe that it should be the responsibility of the university library, ITS Unit, and personal preservation efforts. Some chemistry researchers in FUTA want the library and ITS unit to be responsible for this long-term preservation while other chemistry researchers will prefer to use cloud storage services for preservation or adopt personal preservation e.g. locking hard copies of data files in an iron box or cabinet. FUTMIN chemistry researchers would want the library, copyright organisation, ITS unit, bank, cloud storage, regulatory body' offices, as places for their research data to be stored and preserved for long term. Only one researcher opined that no place is safe for data preservation. Majority of chemistry researchers in FUTO said the library is the place that should be responsible for long term preservation of data, while others opined that it should be the ITS unit, copyright offices or departmental library. Preservation, according to MAUTECH chemistry researchers, should be done either personally, on the internet using email, by the ITS unit, or library. On the contrary, one of the chemistry researchers is of the opinion that self-preservation is adequate for research. A chemistry researcher from FUTO observed that:

> *"…anyway, most of my data I preserve it myself"*

Another researcher however opined that:

> *"I will also what my work to be preserved in the library so that somebody in future can look at my work and see if am wrong or can build his own work".*

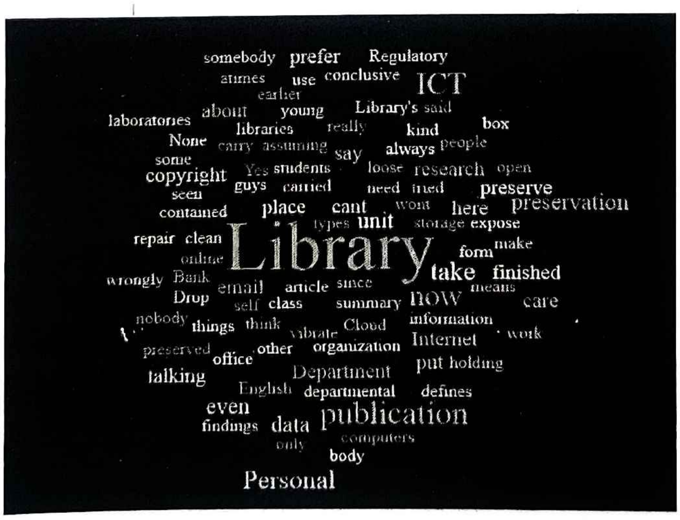

Figure 9: Perception of Chemistry researchers on Where to Preserve Data

6. Data Management Plan (DMP)

Surprisingly, very few chemistry researchers in FUT in Nigeria were aware of a data management plan and as such, they were unable to give a concrete definition of DMP. None of the chemistry researchers in ATBU was aware of a data management plan. They were however willing to learn and be guided on how to manage their data for long-term preservation and reuse. None of FUTA chemistry researchers was aware of a data management plan; however, they see it as a good research practice that should be adopted by all researchers across disciplines. They also require awareness and enlightenment on the need for data management plan, data safety, as well as global best practices on data management. None of the researchers in FUTMIN has heard about a data management plan (DMP). However, researchers indicated their interest in learning how to store, preserve and secure research data. They also need guidance on how to use the repository and any other awareness and guidance available, preferably a computer-based research data management. Chemistry researchers in MAUTECH are not aware of a DMP. They are however, willing and ready to learn and adopt any available and relevant research data management system, especially a computer-based data management system. Here, a chemistry researcher shared a story that:

> *"I know of a case, international case, where a Nigerian did a work and somebody took the work. When the Nigerian saw it, he complained to the body and it was an international case. It was resolved when they ask those guys that claimed the result to bring the evidence of the raw data so it was the raw data that vindicated the Nigerian who has been keeping record…".*

Few chemistry researchers in FUTO were aware of a data management plan (DMP). These chemistry researchers believe that a DMP makes research easier, helps to monitor online visits, saves costs, eliminates boredom and fatigue and overall, makes life easier. Requirement from funders, writing a paper for impact factor journals like *Nature* and *American Chemical Society*, and avoiding excessive spending are noted as some of the motivations for writing a DMP.

Discussion of Findings

Chemistry researchers in federal universities of technology in Nigeria have a good understanding of research data and they generate and use lots of experimental data among the other types of data listed. These findings are in line with Chen & Wu (2017) where 79.83% researchers generate experimental data and also in Aydinoglu et al. (2017) where 52.8% use experimental data. Chemistry researchers experience on data loss was supported by Gonzalez & Peres-Neto (2015) where they observed that "information is lost when researchers fail to store, archive or share their data, this could be as a result of ageing technology or corruption of data-storage devices''. Smith (2014) also observed that where research data is stored is vitally important for current and future use. Improper and/or insecure data storage locations can lead to data loss, theft, unintended deletion, and a plethora of complexities if not stored in stable, secure, and reliable locations. This opinion was also noted by Kuula & Borg (2008) that research data collected often remain in the hands of original researchers without any long-term preservation plan. This may however, impede appropriate access to research data and its reuse for future researches. Research data loss experiences shared by chemistry researchers in Nigeria is not too different with the experiences of their counterparts in the University of Punjab in India where researchers opined for assistance regarding storage and security of data, improving the quality of backup, support for storage and preservation of data (Piracha & Ameen, 2018). Similarly, Buys & Shaw (2015) observed that the types of storage solutions mostly used by chemistry researchers in Nigeria will limit data sharing and long-term preservation. Estimating the expected research data size is also difficult for chemistry researchers hence, establishing a correctly sized research storage system either on the institutional repository or other appropriate storage systems will be difficult (Buys & Shaw, 2015).

Library, as the preferred place for data preservation is supported by Flores et al. (2015) study where it opined that "the library is well situated to be a key player in data management, curation, and preservation, given its extensive experience with selection, metadata, collections, institutional repositories, preservation, curation and access". RDM offers an opportunity for libraries to reformulate their role in the life of the university. Libraries offering RDM services can have a great impact on their campus communities by supporting communication among researchers, enhancing knowledge of the data life cycle, providing disciplinary and institutional resources, and emphasizing the importance of documentation of data sharing (McLure et al. 2014). Similarly, Tenopir et al. (2017) observed that academic library is the major stakeholder in the effective management and preservation of research data. In addition, Tripathi et al. (2017) emphasized on the need for policy formulation by the institution on RDM to encourage and support adherence to RDM as research best practice. Librarians in Nigeria should, as a matter of urgency, leverage on chemistry researchers' willingness to acquire RDM skills and start rolling out relevant research data services that will readily fit into the existing research workflow of researchers. Libraries should also support their staff through training and retraining on RDM skills and knowledge so they can assume their emerging roles as data curators within the scientific community (Tammaro et al., 2018).

Some challenges with the use of a data management plan (DMP) are that it allows for pre-emptive data, narrow the scope of work, and difficulty in complying with different disciplines and funders requirements. This situation was in line with VanTuyl & Michalek (2015) and Şewerin et al. (2014)

study where it was observed that there is the need for harmonization of funding requirements by funding agencies as regards data management plan.

Finally, there is no significant difference in RDM practices among chemistry researchers across federal universities of technology in Nigeria (see Appendix B). However, from the interactions with chemistry researchers during the interview sessions, early-career researchers within the Lecturer II-Senior Lecturer ranks shows a better understanding and appreciates the RDM processes than their older colleagues in the rank of Associate Professors-Professors. This slight difference in perceptions and practices was also noted by Borghi & Van Gulick (2018) in their study of perceptions and practices of MRI researchers in neuroimaging.

Conclusion

This study concludes that there is the need for libraries in FUTs in Nigeria to continue to strive to achieve their research support objectives in their host universities. The need for research data management cannot be overemphasized as research data are increasingly been required to be submitted alongside articles for publication, one of the requirements for grant proposals, and as a means of verifying results and replicating researches in future. Research data management awareness and practice of chemistry researchers in federal universities of technology in Nigeria is low. Chemistry researchers however opined that the awareness, advocacy and training should be carried out by the library which they see as the safest and trusted storehouse of intellectual contents from the university.

Recommendations

1. Libraries of federal universities of technology in Nigeria should embark on education of chemistry researchers on the importance and benefits of research data management in their

institutions. This can be through workshop programs, advocacy, and seminars on research data management.

2. The university management should formulate policies that would encourage submission of underlying datasets along with articles submitted to the institutional repositories. There should be periodic evaluation of such submissions to ascertain that researchers are complying with the directives.

References

Abduldayan, F. J., Dang, T. L., Karemani, A., & Obadia, S. B. (2016). "The role of academic libraries in enhancing workflow in African universities". In *CEUR Workshop Proceedings*, 1830, 158–163. CEUR-WS.

Abduldayan, F.J. (2020). "Research data management: The Nigerian perspectives", Mendeley Data, V2, doi: 10.17632/74ckghv5fn.2, http://dx.doi.org/10.17632/74ckghv5fn.2

Aydinoglu, A.U., Dogan, G. and Taskin, Z. (2017). "Research data management in Turkey: perceptions and practices". *Library Hi Tech*, 35(2), 271-289. https://doi.org/10.1108/LHT-11-2016-0134.

Borghi J.A. and Van Gulick A.E. (2018). "Data management and sharing in neuroimaging: Practices and perceptions of MRI researchers". PLoS ONE 13(7): e0200562. https://doi.org/10.1371/journal.pone.0200562

Braun, V. and Clarke, V. (2006). "Using thematic analysis in Psychology". *Qualitative Research in Psychology,* 3(2), 77-101pp. ISSN 1478-0887, retrieved from: http://eprints.uwe.ac.uk/11735

Buys, C. M. and Shaw, P. L. (2015). "Data management practices across an institution: survey and report". *Journal of Librarianship and Scholarly Communication*, 3(2), eP1225, http://dx.doi.org/10.7710/2162-3309.1225

Chen, X. and Wu, M. (2017). "Survey on the needs for chemistry research data management and sharing". *The Journal of Academic Librarianship,* 43, 346–353. http://dx.doi.org/10.1016/j.acalib.2017.06.006

Chiware, E. and Mathe, Z. (2015). "Academic libraries' role in research data management services: a South African perspective". *South Africa Journal of Libraries and Information Science*, 81(2), 10pp.

Cox, A.M., Kennan, M.A., Lyon, L, and Pinfield, S. (2017). "Developments in research data management in academic libraries: Towards an understanding of research data service maturity". *Journal of the Association for Information Science and Technology*, 68(9):2182–2200.

Dudovskiy, J. (2005). "Research methodology". Retrieved from https://research-methodology.net/sampling-in-primary-data-collection/stratified-sampling/

Flores, J.R., Jason, J.B., Morgan, G.D., Natsuko, N. and Ece, T. (2015). "Libraries and the research data management landscape, the process of discovery". The CLIR Postdoctoral Fellowship Program and the Future of the Academy, 82–102, retrieved from www.clir.org/pubs/reports/pub167/pub167.pdf

Gonzalez A., and Peres-Neto, P.R. (2015). "Act to staunch loss of research data". *Nature correspondence,* Retrieved from www.nature.com/nature/journal/v520/n7548/pdf/520436c.pdf

Helmenstine, A. M. (2016). "Why study chemistry?" Retrieved from www.chemistry.about.com/od/studyskills/f/why-study-chemistry.html

Jones, S., Guy, M, and Pickton, M. (2013). "Research data management for libraries". Retrieved from www.dcc.ac.uk/sites/default/files/documents)events/RDM-for-librarians/RDM-for-librarians-booklet.pdf

Kuula, A., and Borg, S. (2008). "Open access to and reuse of research data: the state of the art in Finland". Retrieved from www.fsd.uta.fi/fi/julkaisut/julkaisusarja/FSDjs07_OECD

McLure, M., Allison, V.L., Catherine, L.C. and Beth, O. (2014). "Data curation: A study of researcher practices and needs". *Portal: Libraries and the Academy,* 14 (2): 139–164. Retrieved from https://muse.jhu.edu/login?auth=0&type=summary&url=/journals/portal_libraries_and_the_academy/v014/14.2.mclure.pdf

Piracha, H.A. and Ameen, K. (2018). "Research data management practices of faculty members". *Pakistan Journal of Information Management & Libraries* (PJIM&L), 20, 60-75.

Provalis QDA Miner (2018). "Qualitative data analysis software (version 5)". Retrieved from www.provalisresearch.com

Raju, R. and Schombee, L. (2013). "Research support through the lens of transformation in academic libraries with reference to the case of Stellenbosch university libraries". *South African Journal of Library and Information Science,* Retrieved From http://sajlis.journals.ac.za DOI: http://dx.doi.org/10.7553/79-2-155

Sewerin, C., Dearborn, D., Henshilwood, A., Spence, M., and Zahradnik, T. (2014). "Research data management faculty practices: A Canadian perspective". *Proceedings of the IATUL Conferences.* Paper 2. Retrieved from http://docs.lib.purdue.edu/iatul/2015/mrd/2, University of Toronto Libraries, Canada.

Smith, P.L. (2014). "Exploring the data management and curation practices of scientists in research labs within a research university". (Doctoral Dissertation of the Florida State University, College of Communication and Information; USA). Retrieved from platosmith.com/fsuphd.14/2014_summer_dissertation.pdf

Tammaro, A.M., Matusiak, K.K., Sposito, F.A., Casarosa, V. (2018). "Data curator's roles and responsibilities: An international perspective". *LIBRI*, 69(2): 89–104. https://doi.org/10.1515/libri-2018-0090

Tenopir C, Allard S, Douglass K, Aydinoglu A.U., Wu L. (2011). "Data sharing by scientists: practices and perceptions". *PLoS ONE*, 6(6), doi:10.1371/journal.pone.0021101.

Tripathi, M., Chand, M., Sonkar, S.K., Jeevan, VKJ (2017). "A brief assessment of researchers' perceptions towards research data in India". *IFLA Journal*, 1-18, DOI: 10.1177/0340035216686984

Van Tuyl, S. and Michalek, G. (2015). "Assessing research data management practices of faculty at Carnegie Mellon University". *Journal of Librarianship and Scholarly Communication,* 3(3). http://dx.doi.org/10.7

APPENDIX A

INTERVIEW SCHEDULE FOR CHEMISTRY RESEARCHERS IN FEDERAL UNIVERSITIES OF TECHNOLOGY IN NIGERIA

Date:
Start:
Stop:
University:

The researcher is interested in understanding your data management practices including how you manage your research data, and the processes, methods and tools used when working with data during your research process.

<u>**Section One**</u>

In this section, the researcher wants to know how you create and manage data and files, the techniques you have developed to organize, format and record the details of your research data.

1. What do you understand by the term 'data'?
2. What do you understand by research data?
3. What are the various types of research data you know?
4. Do you generate digital data?
5. What are the different types of digital data generated? E.g. spreadsheets, texts, images, database, video, Matlab, scanned documents, NMR spectroscopy, etc.
6. Can you name some of the software used in digital data generation?
7. How do you name your stored data?
8. How do you manage different versions of research data generated during experiments?
9. How do you record details of research data e.g. name, source, date etc.
10. Have you ever experienced any situation where you have lost data due to poor file storage, unrecognized file format or bad storage device?
    a. Can you describe what happened?
    b. Did you find a solution? What did you learn from it?

<u>**Section Two**</u>

In this section, the researcher wants to know how you decide your storage needs and manage storage and back up.

1. How do you decide which storage media to use for your data files?
2. What factor do you consider in choosing a storage media for storing your research data?
3. Do you keep record of the media and location of files anywhere?
4. Are you aware of the advantages and disadvantages of different storage media?
5. How do you estimate size of research data?
   a. Is this easy or difficult to do?
6. Can you estimate the likely growth of files?
7. What media do you use for additional storage?

8. How do you protect or secure your research data?
9. Can you explain your back up methods?
   a. Do you record details of back up files? e.g. location, date of back up?
10. How would you describe a situation where you had hardware or software failures and inability to access old files?
   a. What did you or would you do?
11. Would you like advice and guidance on data storage and back up?

Section Three

In this section, the researcher wants to know how you preserve and reuse your research data

1. What would you consider to be long term preservation – in years?
2. When considering preserving data for long term, how would you decide?
   a. What data to preserve for long term?
   b. How long to preserve the data?
3. Where would you go for help in case of research data preservation and related problems?
4. Would you like additional advice and guidance on research data preservation or reuse?

Section Four

In this section, the researcher wants to know the plans and policies which may influence your data management.

1. Are you aware of a Data Management Plan (DMP)?
   a. What is a data management plan?
2. What could motivate you to develop a data management plan?
3. What benefits of developing a data management plan are you aware of?
4. What are the possible challenges associated with developing a data management plan?

Thank you for your time!

APPENDIX B

DIFFERENCES AND SIMILARITIES IN OPINIONS OF CHEMISTRY RESEARCHERS ON RESEARCH DATA MANAGEMENT (RDM) IN FEDERAL UNIVERSITIES OF TECHNOLOGY IN NIGERIA

	FUTA	MAUTECH	ATBU	FUTMX	FUTO	SIMILARITIES
SECTION ONE QUESTIONS						
Q1	Data is a collection of facts gathered from experiments	Valuable information, collected for further study	Data are sequence of numbers or figures derived from experiment for validation	Results gathered from investigations	Quantitative facts gathered from experiment used for further research	Data are outcomes of experiments or investigations emanating from research for further studies
Q2	Collection of research data from laboratories	Results gotten from experiments for finding out things	Data from research work	Facts acquired from research	Results derived from systematic study	Results gathered from systematic studies
Q3	Laboratory data, experimental data, primary data	Experimental Data, field Data	Experimental data, field data, observation, physical and chemical data	Experimental data, primary, field, statistical data	Laboratory data, experimental data, figures	Experimental, laboratory and field data
Q4	Yes, processed data	Yes	Yes, based on the research being carried out	Yes	Yes	Yes
Q5	Spreadsheets, texts, images, NMR spectroscopy, graphs	spreadsheets, images, graphs, analysis	NMR, graphs, image, matlab	Spreadsheets, text, image, video, Matlab, NMR	Graphs, tables, text, images, spreadsheets	Graphs, tables, text, images, spreadsheets, NMR, graph
Q6	Excel, SPSS, NMR spectrometry	Chemdraw, SPSS	Minitab, SPSS, Origin 8, Origin 10, chemdo, material studio, acces	Minitab, SPSS, Excel, chemdraw,	Origin, Excel, Spss	Origin, Excel, SPSS
Q7	Research dependent, self devised, sample name	Self devised	source, name, title, codes, date	Coding with names of experiments, numbers, species	Research based, name	Research dependent, Source, title, date
Q8	Use of defined abbreviations, such as codes, dates	Self devised, using initials	Use NMR, UV, codes, Numbers	Date, name, code	Numbers, dates	use of codes, including Numbers and Alphabets
Q9	name, source, date	Acronyms, codes such as numbers	the computer generates it	Date, name, source	Name, source, date	Name, source and date, atimes generated by computer
Q10	Yes	Yes	Yes, it happens	Yes	Yes	Yes
Q10a	Painful experience at loss of data	System malfunctions, theft	Theft, system, malfunction, corrupt file	System failure, Virus, Mechanical fault, bad instrument	External storage, damage of important data lost	Loss of data due to system failure, theft
Q10b	Backup files and have data other places	Backup using external storage provides security	External disk	Backup, meet specialist, and safeguard data in many places	Backing it up	Back up and look for specialist
SECTION TWO QUESTIONS						
Q1	Portability, accessibility, security	depends on suitability and ease of work	Depends on the value of work on is doing	Reliability, Availability, External hard disk, portability, online storage	Durability, storage capacity, portability	Depending on the kind of work done; Durability, Reliability, Availability, portability
Q2	Safety, reliability, durability	Ease of accessibility, Virus free, and compatibility	Security using external devices, google drive	Security, volume, accessibility	Safety of data, durability, Mobility, portability	Safety, accessibility, durability, Mobility, portability, security
Q3	No	Yes	Save data in record as files	Yes, system based	personally in the office	Yes, but some prefer to save it personally
Q4	YES	Yes	Yes I am aware	Yes	Yes	Yes
Q5	Megabytes, gigabyte and terabyte	No, It depends	It depends, ranges from megabytes to Gigabytes	Megabyte	Gigabyte	It depends on the research but ranges from Megabyte to terabyte
Q5a	No	Not easy		No	No	No
Q6	No	No		No	yes, atimes in megabytes, but atimes its too large	No
Q7	External storage	External storage, hard copy	Google drive, cloud storage, flash CDs, other external storage	Physical and online external storage, CDs, drop box, flash drives	External storage, cloud, emails, flash, Disks ,	On the internet, External storage, flash, Discs etc
Q8	Security using password	Password	Password, CDs, physical cabinet	Password, continous scanning, mails, and cds	Password	Password and other physical security, like locking shelf
Q9	Updates	External Hard disk	Update them regularly	Updates, external storage	Regular updates	Regular Updates
Q9a	Yes	Different locations		Flash drives		Save in different locations
Q10	Yes	Yes	Inaccessibility of files from mail, system crash	Yes	Yes, very frustrating	Very sad experience, due to system failure
Q10a	Seek assistance in resolving problem	Proper Backup to avert forever Data loss		Data loss, seek support externally, and update files	Lost data and seek support for restoring data	loss of data, but seeked support
Q11	Yes	Yes	Share file and save on the cloud	Yes	Yes	Yes
SECTION THREE QUESTIONS						
Q1	For a long period of time	As long as possible, forever	As long as possible	As long as possible, very long time	As long as possible	For a long period of time
Q2	Self Preservation				All data, especially analysed	preserve almost data personally
Q2a	A part or all the data	All data, processed	Some aspect of the data	The useful part of the data	Everything	Processed data
Q2b	As long as its published	For long period of time			As long as possible	As long as possible
Q3	ICT unit, Library	Internet and Library	Library, ICT unit, and personal preservation	Library, ICT, Secured places and organisations	Library	Library and internet based organisations
Q4	Yes	Yes	definitely	Yes	Yes	Yes
SECTION FOUR QUESTIONS						
Q1	No	No	No	No	Not really	No
Q1a						
Q2	Good methods of doing research		We will appreciate it		Good funding to aid article writing	Researchers will appreciate good methods of doing research
Q3	Sharing and preserving to ease life in future		It will have benefit, when it has been used		Yes, it makes life better	Sharing data will make life better
Q4	Creating awareness on good storage media		Storage problem		It will aid research in various disciplines	Create awareness so that it can aid various research in different disciplines

FUTA - FEDERAL UNIVERSITY OF TECHNOLOGY AKURE
MAUTECH - MODIBBO ADAMA UNIVERSITY OF TECHNOLOGY YOLA
ATBU - ABUBAKAR TAFAWA BALEWA UNIVERSITY BAUCHI
FUTMX - FEDERAL UNIVERSITY OF TECHNOLOGY MINNA
FUTO - FEDERAL UNIVERSITY OF TECHNOLOGY OWERRI

The summarised data transcription of all respondents has been deposited on the Mendeley Data Repository. Preview of the dataset is available on this link:

https://data.mendeley.com/datasets/74ckghv5fn/2