

**Data Preservation, Sharing and Reuse Practises
among Agricultural Researchers in Tanzania**

Nolasko Victory Mwinami

**A Thesis Submitted in Fulfilment of the Requirement for the Degree of Doctor of
Philosophy of Sokoine University of Agriculture, Morogoro, Tanzania.**

EXTENDED ABSTRACT

Agricultural research data preservation, sharing, and use practices are crucial in enhancing research activities in Tanzania. However, little is known about the preservation, sharing, and use/reuse practises among agricultural researchers in the country. The main purpose of this study was to investigate data preservation, sharing, and use/reuse practises among agricultural researchers in Tanzania. Specifically, the study, focused on data sources for accessibility and reuse practises among agricultural researchers in Tanzania; analysed the communication channels and their potential applicability in enhancing agricultural research data sharing among agricultural researchers in Tanzania; examined the data preservation practises for enhancing agricultural research data usage among researchers in Tanzania, and examined the factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania.

The current study employed cross-sectional survey methods to investigate data preservation, data-sharing, and use/reuse practises among agricultural researchers in Tanzania. The study was conducted in Agricultural Research Institutions (TARI) centres and Sokoine University of Agriculture (SUA) as an academic and agricultural research institution. The study used both quantitative and qualitative approaches in data collection and analysis. The study involved 227 agricultural researchers who were randomly selected from the study area. Purposively, eleven (11) key informants were selected, namely eleven Directors/Managers from selected TARI centres and one (1) Director of Postgraduate Studies, Research, Technology Transfer & Consultancy for SUA. Also, two Focus Group Discussions which included six participants from Mikocheni TARI centre and Ilonga TARI centre) were conducted in these two centres. The data collection involved the use of multiple data-gathering techniques (questionnaires and interviews and focus group discussions). Quantitative data were analysed with the help of the Statistical Product Service Solution (SPSS) Version 22 while thematic analysis was used to analyse qualitative data. Descriptive statistics techniques were calculated to establish the demographic characteristics of respondents, data sources and use/reuse practises, existing communication channels and their applicability in data sharing, data preservation practices in enhancing data usage, and factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania.

The findings indicate that agricultural researchers use/reuse data that are accessed from different other sources. Findings reveal that the majority of researchers 95.1% use/reuse data in writing scientific papers. Research data are also used in writing technical reports and project proposals. Similarly, results from Binary Logistic Regression Analysis revealed that factors influencing researchers in the use/reuse of research data were not statistical significance meaning that all factors (independent variables) contributed equally to influence researchers in actual data use/reuse.

Furthermore, findings reveal that both mediated and non-mediated channels exist and were used as data-sharing channels. The majority of researchers more than 80% preferred to use non-mediated channels. Moreover, findings show that some mediated data-sharing channels were least used, including social media platforms, institutional repositories and funding agency databases. Through cross-tabulation, findings indicate that the extent of data sharing channel usage at SUA was higher than at the other

research institutions investigated. Some channel usage also has statistical significance including meetings and conferences $p=0.005$, website $p=0.001$, workshop $p=0.049$, telephone as interpersonal mediated $p=0.045$, and technical report $p=0.045$. This meant that these channels were preferred for use across all agricultural research institutions. Findings also indicate that more than fifty percent of the respondents agreed that timely delivery, the cost for the use of a channel and convenience to a channel were among the factors influencing researchers in their channel selection. Likewise, findings reveal that the majority of researchers preferred to preserve their data using different storage devices such as field notebooks, desktop/personal computers, and institutional libraries.

Moreover, the findings indicate that agricultural researchers preferred to preserve their data for more than six years after the end of the project. The findings also indicate the factors that influence researchers in the choice of data preservation methods were: easy to reach, cost-effective storage devices, support to use the devices, adequate infrastructure for data preservation, and reliable power supply. Furthermore, the findings indicate that some factors that influence researchers in data sharing indicating statistical significance included perceived community benefit, perceived ability to share, data sharing norms, and funding agency policies. The findings also indicate that the lack of data sharing policy and the absence of data preservation/management infrastructures, for example, data repositories were the challenges facing agricultural researchers in data sharing.

Based on these findings it can be concluded that agricultural researchers share their data with fellow close research fellows. Agricultural researchers use both mediated and non-mediated data-sharing channels. The main factors motivating researchers to share their data include perceived community benefits, perceived ability to share, data-sharing norms and funding agency policy. Also, it can be concluded that there is yet a great role of research data preservation in enhancing data usage among researchers in Tanzania.

Keywords: Agricultural research data, data sharing, data sharing channels, data reuse, research data.

MUHTASARI ULIOPANULIWA

Uhifadhi wa data za utafiti wa kilimo, ushiriki na utumiaji tena ni muhimu katika kuimarisha shughuli za utafiti nchini Tanzania. Hata hivyo, kidogo inajulikana kuhusu kuhifadhi, kushiriki, na kutumia tena desturi za watafiti wa kilimo nchini. Madhumuni makuu ya utafiti huu yalikuwa ni kuchunguza uhifadhi, ushirikishwaji na utumiaji tena wa data miongoni mwa watafiti wa kilimo nchini Tanzania. Hasa, utafiti ulizingatia vyanzo vya data kwa ajili ya upatikanaji na utumiaji tena wa data kwa watafiti wa kilimo nchini Tanzania; kuchambua njia za mawasiliano na uwezekano wake wa kutumika katika kuboresha ugawanaji wa takwimu za utafiti wa kilimo miongoni mwa watafiti wa kilimo nchini Tanzania; ilichunguza mazoea ya kuhifadhi data kwa ajili ya kuimarisha matumizi ya data ya utafiti wa kilimo miongoni mwa watafiti nchini Tanzania, na kuchunguza mambo yanayoathiri mbinu za watafiti wa kilimo wa ugawanaji data ili kuboresha upatikanaji wa data za utafiti wa kilimo nchini Tanzania.

Utafiti wa sasa ulitumia mbinu za uchunguzi wa sehemu mbalimbali kuchunguza uhifadhi wa data, ushiriki wa data na utumiaji tena wa watafiti wa kilimo nchini Tanzania. Utafiti huo ulifanyika katika vituo vya Taasisi za Utafiti wa Kilimo (TARI) na Chuo Kikuu cha Sokoine cha Kilimo (SUA) kama taasisi ya utafiti wa kitaaluma na kilimo. Utafiti ulitumia mbinu za upimaji na ubora katika ukusanyaji na uchambuzi wa data. Utafiti ulihusisha watafiti wa kilimo 227 ambao walichaguliwa bila mpangilio kutoka eneo la utafiti. Kwa makusudi, watoa taarifa muhimu kumi na moja (11) walichaguliwa, ambao ni Wakurugenzi/Mameneja kumi na moja kutoka vituo vilivyochaguliwa vya TARI na Mkurugenzi mmoja (1) wa Mafunzo ya Uzamili, Utafiti, Uhamisho wa Teknolojia na Ushauri wa SUA. Pia, Mijadala miwili ya Vikundi Mwelekeo iliyojumuishwa washiriki sita kutoka kituo cha Mikocheni TARI na kituo cha Ilonga TARI ilifanyika katika vituo hivi viwili. Mkusanyiko wa data ulihusisha matumizi ya mbinu nyingi za kukusanya data (hojaji na mahojiano na mijadala ya vikundi lengwa). Data ya kiasi ilichanganuliwa kwa usaidizi wa Suluhu la Huduma ya Kitakwimu (SPSS) Toleo la 22 huku uchanganuzi wa kimaudhui ukitumika kuchanganua data ya ubora. Mbinu za takwimu za maelezo zilikokotolewa ili kubaini sifa za idadi ya watu waliohojiwa, vyanzo vya data na mazoea ya utumiaji/utumiaji tena, njia zilizopo za mawasiliano na ufaafu wao katika kushiriki data, mbinu za kuhifadhi data katika kuboresha matumizi ya data, na mambo yanayoathiri mazoea ya kushiriki data ya watafiti wa kilimo ili kuimarishwa. upatikanaji wa takwimu za utafiti wa kilimo nchini Tanzania.

Matokeo ya utafiti yanaonyesha kuwa watafiti wa kilimo hutumia/hutumia tena data ambayo hupatikana kutoka kwa vyanzo vingine tofauti. Matokeo yanaonyesha kuwa watafiti wengi 95.1% hutumia tena data katika kuandika karatasi za kisayansi. Data ya utafiti pia hutumiwa katika kuandika ripoti za kiufundi na mapendekezo ya mradi. Vile vile, matokeo ya Uchanganuzi wa Urekebishaji wa Usafirishaji wa Binary yalifichua kuwa mambo yanayoathiri watafiti katika utumiaji/utumiaji tena wa data ya utafiti hayakuwa umuhimu wa kitakwimu ikimaanisha kuwa mambo yote (vigezo vinavyojitegemea) vilichangia kwa usawa kuathiri watafiti katika utumiaji halisi wa data.

Zaidi ya hayo, matokeo yanaonyesha kuwa njia zote mbili zilizopatanishwa, na zisizo za upatanishi zipo na zilitumika kama njia za kushiriki data. Watafiti wengi zaidi ya 80% walipendelea kutumia njia zisizo za upatanishi. Zaidi ya hayo, matokeo yanaonyesha kuwa baadhi ya njia za ushiriki wa data zilizopatanishwa zilitumiwa kwa kiwango cha chini, ikiwa ni pamoja na majukwaa ya mitandao ya kijamii, hazina za kitaasisi, na hifadhidata za

wakala wa ufadhili. Kupitia kwa majedwali mtambuka, matokeo yanaonyesha kuwa kiwango cha utumiaji wa chaneli ya kushiriki data katika SUA kilikuwa cha juu kuliko katika taasisi nyingine za utafiti zilizochunguzwa. Baadhi ya matumizi ya chaneli pia yana umuhimu wa kitakwimu ikijumuisha mikutano na makongamano $p=0.005$, tovuti $p=0.001$, warsha $p=0.049$, simu kama upatanishi baina ya watu $p=0.045$, na ripoti ya kiufundi $p=0.045$. Hii ilimaanisha kuwa njia hizi zilipendelewa kutumika katika taasisi zote za utafiti wa kilimo. Matokeo pia yanaonyesha kuwa zaidi ya asilimia hamsini ya waliojibu walikubali kuwa uwasilishaji kwa wakati, gharama ya matumizi ya chaneli na urahisi wa kituo ni miongoni mwa mambo yanayoathiri watafiti katika uteuzi wa chaneli zao. Vile vile, matokeo yanaonyesha kuwa watafiti wengi walipendelea kuhifadhi data zao kwa kutumia vifaa tofauti vya kuhifadhi kama vile madaftari ya shambani, kompyuta ya mezani/ya kibinafsi, na maktaba za taasisi.

Aidha, matokeo yanaonyesha kuwa watafiti wa kilimo walipendelea kuhifadhi data zao kwa zaidi ya miaka sita baada ya kumalizika kwa mradi huo. Matokeo hayo pia yanaonyesha mambo yanayoathiri watafiti katika uchaguzi wa mbinu za kuhifadhi data ni haya: rahisi kufikiwa, vifaa vya uhifadhi wa gharama nafuu, usaidizi wa kutumia vifaa, miundombinu ya kutosha kwa ajili ya kuhifadhi data, na usambazaji wa umeme unaotegemewa. Zaidi ya hayo, matokeo yanaonyesha kuwa baadhi ya vipengele vinavyoathiri watafiti katika ugavi wa data vinavyoonyesha umuhimu wa takwimu vilijumuisha manufaa ya jumuiya inayotambulika, uwezo unaotambulika wa kushiriki, kanuni za kushiriki data na sera za wakala wa ufadhili. Matokeo ya utafiti pia yanaonyesha kuwa kukosekana kwa sera ya kugawana data na kutokuwepo kwa miundombinu ya kuhifadhi/kusimamia data, kwa mfano, hazina za data zilikuwa changamoto zinazowakabili watafiti wa kilimo katika ugawanaji wa data.

Kulingana na matokeo haya inaweza kuhitimishwa kuwa watafiti wa kilimo hutumia njia zote mbili za upatanishi na zisizo za upatanishi wa kushiriki data. Pia, watafiti wa kilimo hushiriki data zao na wenzao wa utafiti wa karibu. Sababu kuu zinazowahamasisha watafiti kushiriki data zao ni pamoja na faida zinazochukuliwa kuwa za jumuiya, uwezo unaotambulika wa kushiriki, kanuni za kushiriki data na sera ya wakala wa ufadhili. Pia, inaweza kuhitimishwa kuwa bado kuna nafasi kubwa ya kuhifadhi data za utafiti katika kuimarisha matumizi ya data miongoni mwa watafiti nchini Tanzania.

DECLARATION

I, Nolasko Victory Mwinami, do hereby declare to the Senate of Sokoine University of Agriculture that this thesis is my original work done within the period of registration has not been presented and will not be presented nor is concurrently being submitted to any other institution.

Nolasko Victory Mwinami (PhD Candidate)

Date

The above declaration is confirmed by:

Prof. Frankwell W. Dulle
(Supervisor)

Date

Prof. Wulystan Pius Mtega
(Supervisor)

Date

COPYRIGHT

No part of this thesis may be reproduced, stored in any retrieval system, or transmitted in any form or by any means without prior written permission of the author or Sokoine University of Agriculture on that behalf.

ACKNOWLEDGEMENTS

This work is a result of contributions from the power of God and many people. I cannot mention all of them here because they are so many so the list will be too long. From the bottom of my heart, I thank all those whose support, encouragement, advice and love helped me to accomplish this work. Firstly, I would like to express my gratefulness to the Almighty God for his protection, provision, intellect, health and enabling me to pursue my studies. I kindly appreciate my supervisors Prof. Frankwell W. Dulle and Prof. Wulystan Pius Mtega of Sokoine University of Agriculture (SUA) for their endless supervision, critique, suggestions, guidance, advice, support, and encouragement from the beginning of my study to the end. I wish to extend my thanks to Jordan University College for permitting me to pursue my further studies at SUA. I express my appreciation to Tanzania Agricultural Research Institute (TARI) headquarters for allowing me to conduct my study in TARI centres. I am also grateful to the Tanzania Commission for Science and Technology (COSTECH) for allowing me to attend the scientific paper writing training that added value to my skills and knowledge in writing this scientific paper. Moreover, I would like to thank Dr. Andrew Malekani, Dr. Mashauri Chaila, and Prof. Joseph Maghway for their advice, kindness, and their insightful. Many thanks are also directed to all agricultural researchers at SUA and TARI centres for their cooperation during the data collection process. This work was accomplished with the help of many people who have added their input in one way or another on making me finish this work. Lastly, I wish to thank my family members who advised, encouraged and supported me directly and indirectly.

DEDICATION

This work is dedicated to my late beloved Father Victory Mwinami and Mother Avelina Myonga who passed away many years ago before the dreams of this work were realised. Also, the work is also dedicated to my family members including my wife Emilia Lucian George and children Irene, Victory-Hans, Stephano, and Isaack for their tireless support and understanding of my long absence from home.

TABLE OF CONTENTS

EXTENDED ABSTRACT	ii
MUHTASARI ULIOPANULIWA	iv
DECLARATION	vi
COPYRIGHT	vii
ACKNOWLEDGEMENTS	vi
DEDICATION	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	x
LIST OF FIGURES	xii
LIST OF APPENDICES	xiii
LIST OF ABBREVIATIONS AND ACRONYMS	xiv
ORGANIZATION OF THE THESIS	xvi
CHAPTER ONE	1
1.0 General Introduction	1
1.1 Background to the Study	1
1.2 Problem Statement	4
1.3 Justification of the Study	4
1.4 General objectives	4
1.4.1 Specific Objectives	4
1.4.2 Research Questions	5
1.5 Theoretical and Conceptual Framework of the Study	5
1.5.1 Data lifecycle model for data preservation	5
1.5.2 The Theory of Reasoned Action (TRA)	5
1.5.3 Concentric layered model	5
1.5.4 Theory of planned behaviour	5
1.6 General Methodology	7
1.6.1 Population and sampling	7
1.6.2 Data collection	8
1.6.3 Statistical data analysis	8
1.6.4 Measurement of construct	9
CHAPTER TWO	15
PAPER I	15
CHAPTER THREE	29
PAPER II	29
CHAPTER FOUR	53
PAPER III	53
CHAPTER FIVE	81
PAPER IV	81
Factors influencing data sharing practises, perceptions and motivations for improved accessibility of agricultural research data in Tanzania	81

Abstract.....	81
1.0 Introduction	83
1.1 Statement of the Problem	84
1.2 General Objective.....	85
1.2.1 Specific objectives.....	85
1.2.2 Research questions.....	85
1.2.3 Significance of the study	85
2.0 Literature Review	86
2.1 Data Sharing Practices by Agricultural Researchers.....	86
2.2 Factors influencing data sharing behaviours among researchers	87
2.3 Theoretical Framework.....	88
3.0 Research Method	88
3.1 Population and Sampling.....	89
3.2 Data Collection	89
3.3 Statistical Data Analysis	89
3.4 Measurement of Construct	90
4.0 Findings and Discussion.....	90
4.1 Demographic Characteristics of Respondents.....	90
4.2 Data Sharing among Agricultural Researchers.....	91
4.3 Data sharing stakeholders exhibited by agricultural researchers in the study area	92
4.4 Factors influencing data sharing practice among researchers	93
4.5 Challenges Hindering Data Sharing among Agricultural Researchers	96
5.0 Conclusion	97
6.0 Recommendations	97
References.....	99
CHAPTER SIX	105
6.0 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	105
6.1 Summary of Major Results and Conclusions	105
6.1.1 Data preservation practices for enhancing agricultural research data usage among agricultural researchers	105
6.1.2 Data sources for accessibility and use/reuse practices among agricultural researchers in the study area.....	105
6.1.3 Communication channels and their potential applicability in enhancing agricultural research data sharing among agricultural researchers.....	106
6.1.4 Factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania.	106
6.2 Conclusion	107
6.3 Recommendations	107
6.4 Contribution of the Study to the Body of Knowledge.....	108
6.5 Suggested Areas for Further Research	108
APPENDICES.....	110

LIST OF TABLES

CHAPTER TWO

Table 1:	Sampling frame.....	21
Table 2:	Demographic characteristics of respondents.....	22
Table 3:	Researchers academic field of specialization.....	22
Table 4:	Data preservation methods existing and used in research institutions.....	23
Table 5:	Length of research data preservation in research institutions.....	24
Table 6:	Factors that influence the choice of data preservation method/device.....	25

CHAPTER THREE

Table 1:	Sources of agriculture research data used by researchers	39
Table 2:	Inferential statistics for sources of data for reuse across research institutions.....	41
Table 3:	Type of research data reused by researchers.....	42
Table 4:	Rate of researchers reuse of others researchers' data.....	42
Table 5:	How data are used/reused by researchers.....	42
Table 6:	Factors influencing researchers to use/reuse other researchers' data.....	43
Table 7:	Modeling factors influencing researchers to reuse other researchers' data.....	44

CHAPTER FOUR

Table 1:	Channels existing and used in agriculture research institutions.....	68
Table 2:	Extent of channel usage across research institutions as segregated by institution.....	71
Table 3:	Factors that influenced researchers in channel selection.....	72

CHAPTER FIVE

Table 1:	Data sharing stakeholders exhibited by agricultural researchers.....	93
Table 2:	Factors influencing data sharing practice among researchers.....	96
Table 3:	Challenges hindering data sharing among agricultural researchers.....	96

LIST OF FIGURES

CHAPTER ONE

Figure 1: Research Data Sharing Conceptual Model	6
Figure 2: Conceptual framework showing data preservation, data sharing and reuse	7

CHAPTER TWO

Figure 1: Agricultural research data life cycle model.....	18
Figure 2: Agricultural research data preservation conceptual frame work.....	20

CHAPTER THREE

Figure 1: Conceptual model showing factors that influence data use/ reuse behaviour.....	36
Figure 2: Sampling distribution.....	37
Figure 3: Demographic information of respondents.....	38
Figure 4: Academic field of specialization of researchers.....	39

CHAPTER FOUR

Figure 1: Adopted and modified model for data sharing channel selection.....	61
Figure 2: Conceptual framework showing the relationship exist between the variables for this study.....	62
Figure 3: Sampling distribution.....	63
Figure 4: Characteristics of the respondents.....	66
Figure 5: Researchers' Academic Field of Specialization.....	67

CHAPTER FIVE

Figure 1: Characteristics of the respondents.....	91
Figure 2: Showing data sharing among agricultural researchers.....	92

LIST OF APPENDICES

Appendix 1: Questionnaire for Agricultural Researchers110
Appendix 2: Interview Guide.....117
Appendix 3: Focus Group Discussion Guide118
Appendix 4: Research permits letters119

LIST OF ABBREVIATIONS AND ACRONYMS

AGORA	Access to Global Online Research
AGRIS	International System for Agricultural Science and Technology
AGROVOC	Multilingual Agricultural Thesaurus developed by FAO, covers areas such as food, nutrition, agriculture, forestry, fisheries, scientific and common names of animals and plants, environment, biological notions, and techniques of plant cultivation.
AIMS	Agriculture Information Management System
AIOTI	Alliance for Internet of Things Innovation
CESSDA	European Social Science Data Archive
CIAT	Cenro Internationale de Agricultural Tropical
CIMMYT	International Maize and Wheat Improvement Centre
CONAS	College Of National and Applied Science
COSTECH	Tanzania Commission for Science and Technology
CTA	Technical Centre for Agriculture and Rural Cooperation
DCC	Data Curation Centre
DIIT	Department of Information and Information Technology
DOAJ	Directory of Open Access Journal
DPRTC	Directorate of Postgraduate Studies, Research, Technology Transfer and Consultancy
EUBON	European Biodiversity Network
FAIR	Findable, Accessible, Interoperable, and Reusable
FAO	Food and Agriculture Organization of the United Nations
FAOSTAT	Food and Agriculture Organization Corporate Statistical Database
FGD	Focus Group Discussion
GIS	Geographic Information System
GODAN	Global Open Data for Agriculture and Nutrition
HMIS	Health Management Information System
IBP	International Breeding Platform
ICRISAT	International Crop Research Institute for Semi-arid Tropics
ICTs	Information and Communication Technologies
IITA	International Institute of Tropical Agriculture
IRB	Institutional Review Board
IRI	International Research Institute for Climate & Society
NARS	National Agriculture Research System
NBS	National Bureau of Statistics
NIH	National Institute of Health
NSF	National Science Foundation
NWOW	New Way of Working
OECD	Organization of Economic Cooperation and Development
RDA	Research Data Alliance
RDM	Research Data Management
SPSS	Statistical Package of Social Science
SS	Sample Size
STEM	Science, Technology, Engineering and Mathematics
SUA	Sokoine University of Agriculture
TARI	Tanzania Agriculture Research Institute

TMA	Tanzania Meteorological Agency
TPB	Theory of Planned Behaviour
TRA	Theory of Reasoned Action
UKDA	UK Data Archive Portals
URT	United Republic of Tanzania
USAID	United States Agency for International Development
WHO	World Health Organization

ORGANIZATION OF THE THESIS

This thesis is organised into six chapters. With the exception of Chapter One and Six, all other chapters are organised in the form of a publishable manuscript. Chapter One is a general introduction, problem statement, justification of the study, study objectives, the theoretical and conceptual frameworks of the study, and methodology. Chapter Two present manuscript paper one, Data Preservation Practises for Enhancing Agricultural Research Data Usage among Agriculture Researchers in Tanzania. Chapter Three presents manuscript paper two, a Study of Data Sources for Accessibility and Reuse Practices among Agricultural Researchers in Tanzania. Chapter Four is paper three manuscript about Communication Channels and their Potential Applicability in Enhancing Agricultural Research Data Sharing among Agricultural Researchers in Tanzania.

Chapter Five presents objective four and is manuscript about, factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania. Finally, Chapter six presents a summary of the thesis together with a conclusion and overall recommendations based on the study findings. The status of all papers is as follows;

Paper I

Mwinami, L., Dulle, F. W & Mtega, W.P. (2022). Data preservation practices for enhancing agricultural research data usage among agricultural researchers in Tanzania. *Journal of Librarianship & Information Science*, 1-3. Doi10.1177/09610006221138110.

Paper II

Mwinami, L., Dulle, F. W & Mtega, W.P. (2023). A Study of Data Sources for Accessibility and Reuse Practices among Agricultural Researchers in Tanzania, *Journal of Agricultural & Food Information*, DOI:10.1080/10496505.2023.2190899

Paper III

Mwinami, L., Dulle, F. W & Mtega, W.P. (In-process).Communication Channels and their Potential Applicability in Enhancing Agricultural Research Data Sharing among Agricultural Researchers in Tanzania. *LIBER Quarterly: The Journal of the Association of European Research Libraries*. Vol. 33, 1–28.

Paper IV

Mwinami, L., Dulle, F. W & Mtega, W.P. (Accepted in process for publication). Factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania. *HSO Journal of Environmental Science Current Research*. 6:045, DOI: 10.24966/ESCR-5020/100045

CHAPTER ONE

1.0 General Introduction

1.1 Background to the Study

Research data are data generated through the research process (Tripathi, Skukla and Sonker, 2017). Agricultural research data are the basic data generated through agricultural scientific and technological research and can be termed as the output of the research process (Zhao and Wang, 2015). Agricultural research data comprises remote sensing data, geographic information systems data (GIS), baseline survey data, genomic data, and crop breeding data (Kirub, 2016a). Data can be found in different forms. Data may exist as text documents, graphics, databases, photographs, test responses, spread sheets, laboratory notes, films, and audio data (Chigwada *et al.*, 2017; Elsayed, 2018). Agriculture research data are important in many of its aspects, including aspects that deal with production aiming to feed the global community (Moore *et al.*, 2020). Therefore, data are needed to be preserved and shared through different channels in such a way that users can reuse them in the future.

Data preservation ensures the permanent availability of research data. Data preservation includes the process that ensures efficient data storage so that in the end, such data can be easily accessed, browsed, and used for other research projects (Tripathi, Skukla, and Sonker, 2017). Prevention of destruction and loss of data is the primary goal for researchers (Kadir and Yunus, 2017). Preservation of accurate and reliable data is necessary simply because it will permit future access and sharing for the benefit of science (DPRTC, 2019; Tenopir *et al.*, 2011).

Data are preserved through data storage infrastructures that permit the sharing of such data. Data repositories are infrastructures for data preservation and sharing. For example, agricultural researchers always share their data by depositing it into a database and repositories such as the Global Biodiversity Information Faculty (GBIF) (Parr *et al.*, 2017). Therefore, data preservation enhances data sharing and use/reuse (Kamaruddin *et al.*, 2017; Yoon & Kim, 2019). Therefore, sharing data by depositing it in any data preservation infrastructure will guarantee future availability and accessibility for reuse by another researcher.

Scientific data sharing is important to enhance data use/reuse among researchers. Scientific data sharing is a process of allowing the data to be available for use by other researchers by either depositing or sending it to fellow researchers upon request (Kim & Stanton, 2016; Shen, 2017). Therefore, sharing scientific data provides a chance for other researchers to use/reuse already existing data (Pronk, 2019; Sielemann *et al.*, 2020; Tenopir *et al.*, 2015). In enhancing data sharing scientists are advised to abide by the data sharing principle of data communism. This principle was established by Robert Merton (1942/1973), the principle emphasises that all scientists are needed to have free and equal access to scientific data (Zoran, 2017) the principle emphasise that scientists should always give priority to the progress of science over the interest of individual achievement. There are close linkages between data preservation, sharing, and reuse practises. Effective data management/preservation provides the chance for agricultural scientific data access, sharing, and reuse (Schmidt *et al.*, 2022; Williams *et al.*, 2019). In the same vein, data sharing and reuse are interrelated simply because data sharing provides room for data reuse (Aydinoglu *et al.*, 2017; Griffin *et al.*, 2022; Tenopir *et al.*, 2020; Yoon & Kim, 2019).

Over the recent year, data sharing has attracted more attention from researchers, research institutions, governments, funding agencies, and publishers. Research projects funding agencies such as the National Institute of Health (NIH), United States Agency for International Development (USAID), National Science Foundation (NSF), World Data System, Committee on Data Science and Technology, Research Data Alliance, and Bill and Melinda Gates Foundations require or mandate researchers to share data or open up their data to all funded research projects (Adrian *et al.*, 2018; Bezuidenhout and Chakauya, 2018; Kirub, 2016b; Tenopir *et al.*, 2018). Furthermore, for scientific data to be shared and used effectively, among other things, data ought to be properly managed. The appropriate data management infrastructures include repositories, databases, and archives that can enhance data sharing and use/reuse (Late & Kekalainen, 2020).

There are several global initiatives intended to manage data to enhance scientific data sharing and reuse. The establishment of the Wheat Information System (Wheat IS) allows wheat data producers to share their data and at the same time access other researchers' data that are produced from other parts of the world (Sen *et al.*, 2022). Likewise, several initiatives lead to the establishment of centralised data management systems such as North America Data Archives, European Social Science Data Archives (CESSDA), and UK Data Archives Portals (UKDA) these ensure that global scientific data, knowledge is managed and easily accessed when needed by researchers anywhere (Kamaruddin, Kadir and Yunus, 2017). Furthermore, a recent initiative by Global Open Data for Agriculture and Nutrition (GODAN) and its partners focused on convincing governments worldwide to see the importance of opening up their agricultural data. The G-8 leaders created GODAN to allow African countries and the whole world to have access to relevant scientific agricultural data that are found among the G-8 countries (Musker & Schaap, 2022).

In line with this, there is the establishment of the agINFRA an Innovative Integrated Infrastructure initiative that aims to promote scientific agricultural data sharing and reuse among agricultural scientists worldwide (Drakos *et al.*, 2015). Likewise, the establishment of the International System for Agricultural Science and Technology (AGRIS) initiative aimed at making relevant agricultural research data and information globally available to the user who could share and download their agricultural data and information (Celli *et al.*, 2022). These initiatives intended to open up or make data discoverable to the scientific community have shown that data is the key to scientific advancement in agricultural research. The availability of easily accessible quality scientific agricultural data is crucial for driving food production thus contributing to solving the global problem of food insecurity (Ifeanyi-obi & Ibisio, 2020; Kuteesa & Kyotalimye, 2019; Williamson *et al.*, 2021).

In Africa, there are several initiatives made to strengthen data sharing. This has been shown by the African Data Revolution Report of 2018, which reveals that there is the existence of differences among African countries in embracing Open Government Data. The report exposes that Open Data is inadequately recognised in the African context; therefore, the Government is urged to support data sharing that could benefit all sectors including agriculture (Van Bell *et al.*, 2018). Tanzania was not far behind in the global efforts to make data accessible and discoverable to the scientific community. There has been efforts such as the establishment of the Open Data Policy Draft, and the Open Data Portal for preserving and disseminating different types of scientific data (Agrawal, 2017). Furthermore, the establishment of the e-Government Act in 2019, emphasises the need to create primary data, share and make data available across public institutions in Tanzania

(URT, 2019). This Act of 2019 is a general guideline and not specific to agricultural research institutions. Also, the Tanzania Commission for Science and Technology (COSTECH) which have a responsibility to register all publicly funded research projects; and researchers are also required to submit raw data at the end of their projects (COSTECH, 2022). This implies that researchers are sharing scientific data with COSTECH and affiliated research institutions at the end of their projects. In line with this, the research guideline of 2019 for SUA reveals researchers and students are required to share their data and other research materials at the end of the research (DPRTC, 2019). The sharing practises are upon individual or institution's request whereby the consideration for sharing is based upon institutional and national policies and guidelines. However, the Act of 2016 for Tanzania Agricultural Research Institute (TARI) is silent on data sharing; it only spells out matters regarding technology transfer (URT, 2016). In line with this, the Government of Tanzania established a data transfer procedure in 2010 to guide the transfer of data between Tanzania and foreign institutions (Mboera, 2015; Mushi *et al.*, 2020).

Despite the existence of these data-sharing guidelines and initiatives that emphasise open data, data sharing, and exchange in Tanzania, data preservation, sharing, and reuse practises in agricultural research institutions in Tanzania are relatively unknown. Likewise, the relevant literature shows there is the existence of different types of agricultural data that are preserved, accessed, and shared by some public institutions.

For instance, statistical data, metadata, agricultural administrative data, baseline data gathered through sample census, satellite climate data based on crop conditions, crop stage, crop yields, and production statistics based on agriculture and food are stored and shared via different databases at the National Bureau of Statistics, Tanzania Meteorological Agency and at District level (Bhatia *et al.*, 2016; FAO, 2017; Kijazi, 2018; NBS, 2010; TMA, 2020; URT, 2017a). To harness the data properly there is a need first to preserve such data. The findings of a study that was conducted at the University of Dodoma in Tanzania reveal that better data preservation can facilitate data accessibility, sharing, and reuse among different stakeholders (Mushi *et al.*, 2019). To this end, it can be noted that these initiatives and efforts that intended to make data open and accessible to the scientific community were mostly from developed countries by comparison to developing countries like Tanzania. Despite these efforts and initiatives to share data and make data discoverable in developing countries like Tanzania, there is limited documented evidence that exposes data preservation, sharing, and reuse practises particular in agricultural research institutions in Tanzania. The main purpose of this study was to investigate data preservation, sharing, and reuse practises among agricultural researchers in Tanzania.

1.2 Problem Statement

Agricultural research institutions in Tanzania have rich scientific data which are generated through different research projects. Despite the existence of such potential agricultural research data in research institutions particularly in SUA and TARI centres in Tanzania, scientific data preservation practices, data sharing and reuse practises in research institutions are unknown (Katabalwa *et al.*, 2021; Mushi *et al.*, 2019). As a result, this has limited the advancement in data preservation and sharing that could enhance scientific data reuse for research. However, few studies, initiatives and research guidelines that guide research in Tanzania stipulate some aspects of dissemination, sharing, and exchange of agricultural data such as raw data and other research material sharing, statistical data, metadata, and agricultural administrative data (Bhatia *et al.*, 2016; COSTECH, 2022; FAO, 2017; Kijazi, 2018; TMA, 2020; URT, 2017a; URT, 2016). In line with this, other studies have revealed the existence of health data management systems, and health data sharing policies in Tanzania (Mboera, 2015; MUHAS, 2020 URT, 2017b). Therefore, what has been noticed is a significant gap in the subject under study where agricultural research data preservation, sharing, and reuse practises in agricultural research institutions in Tanzania is not well documented. Therefore, this study was conducted to bridge this knowledge gap to investigate data preservation, sharing, and reuse practises among agricultural researchers in Tanzania.

1.3 Justification of the Study

The study will enlighten the agricultural research institutions to strengthen existing and adopt more formal methods for preserving agricultural scientific data. This study's findings can be used by research institutions and the Government as a base in the process to establish data preservation, sharing, and reuse policies. Such policies could guide research institutions in their data preservation, sharing, and usage practises among researchers locally and internationally. Findings will create awareness in agricultural research institutions to see the need for strengthening the department of technology transfer to perform a coordination function for the management, sharing, and reuse practises in research institutions. This study's findings will enlighten agricultural researchers on different sources for accessing agricultural research data. Accessing scientific data from different sources can help in generating new knowledge and innovation. This study has also generated much knowledge that might motivate scientists to be willing to share their data by considering first the benefits of science over the researcher's interests.

1.4 General objectives

The main objective of this study was to investigate data preservation, sharing, and use/reuse practises among agricultural researchers in Tanzania.

1.4.1 Specific Objectives

- i. Examined research data preservation methods for enhanced data usage among agricultural researchers in Tanzania.
- ii. Investigated data sources for accessibility and reuse practises among agricultural researchers in Tanzania.
- iii. Analysed the communication channels and their potential applicability in enhancing agricultural research data sharing among agricultural researchers in Tanzania.
- iv. Examined the factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania.

1.4.2 Research Questions

Based on our research objectives, our research questions are:

- i. Which data preservation methods are used by agricultural researchers?
- ii. Where do agricultural researchers find data to reuse?
- iii. Which communication channels are used to enhance data sharing in Tanzania?
- iv. Which factors influence agricultural researchers in data-sharing practises?

1.5 Theoretical and Conceptual Framework of the Study

1.5.1 Data lifecycle model for data preservation

This study adopted Data Curation Centre (DCC) Lifecycle Model. The data lifecycle model explains that data sharing is a necessary part of the research lifecycle (Tenopir *et al.*, 2015). The data lifecycle model also emphasises the importance of data preservation infrastructure as an important factor influencing effective data management and enhancing data sharing. The model is relevant to this study because it illuminates data preservation as one of the core activities in the data lifecycle. Data preservation or management is a participatory activity that involves researchers, data managers, curators, repository specialists, librarians, and project managers (Pouchard, 2016). The model expresses the value of data preservation thus without data preservation, data will not be available and easily accessed for use/reuse in future research (Ng'eno and Mutula, 2018).

1.5.2 The Theory of Reasoned Action (TRA)

This study adopted the Theory of Reasoned Action to (TRA) explain factors influencing agriculture researchers on their actual behaviour of data reuse. The theory states that the person's intention is determined by attitudes and subjective norms that can lead to performing actual behaviour (Ajzen and Fishbein, 1980). The theory was relevant to this study because data reuse behaviour is determined by attitudes, subjective norms and the availability of facilitating conditions.

1.5.3 Concentric Layered Model

This study adopted the concentric layered model to guide researchers in factors that need to be considered in the choices of channels for data sharing. The model consists of four layers and each layer consists of factors influencing choices in channel selection (Kok *et al.*, 2016). This model was modified to include factors that could better guide the selection of the data sharing channel.

1.5.4 Theory of planned behaviour

This study adopted the Theory of Planned Behaviour to explain factors that influence researchers in their data-sharing practises. The TPB is a theory that explains people's behaviour as a result of attitudes, and the availability of resources facilitating conditions (Zenk-Möltgen *et al.*, 2018). The Theory of Planned Behaviour is relevant to this study because it guides us in the understanding of individual and institutional factors that can lead to the actual behaviour of data sharing. Several previous studies have reported using this theory in explaining the contribution of individual and institutional factors in data sharing (Yoon & Kim, 2019; Zenk-Moltgen *et al.*, 2018). Figure 1 below shows the data-sharing conceptual model from the Theory of Planned Behaviour (TPB). The model shows the relationship between independent variables (individual factors and institutional factors), and dependent variables (data sharing). Independent variables from TPB comprise individual factors (perceived community benefits, perceived career benefits,

perceived ability to share, scholarly altruism, and data-sharing norms). The institutional factors (funding agency policies, journal publisher's policies, and facilitating conditions). The presence of the individual, institutional, and intervening variable (behavioural intention to share research data) leads to actual data-sharing behaviour (dependent variable).

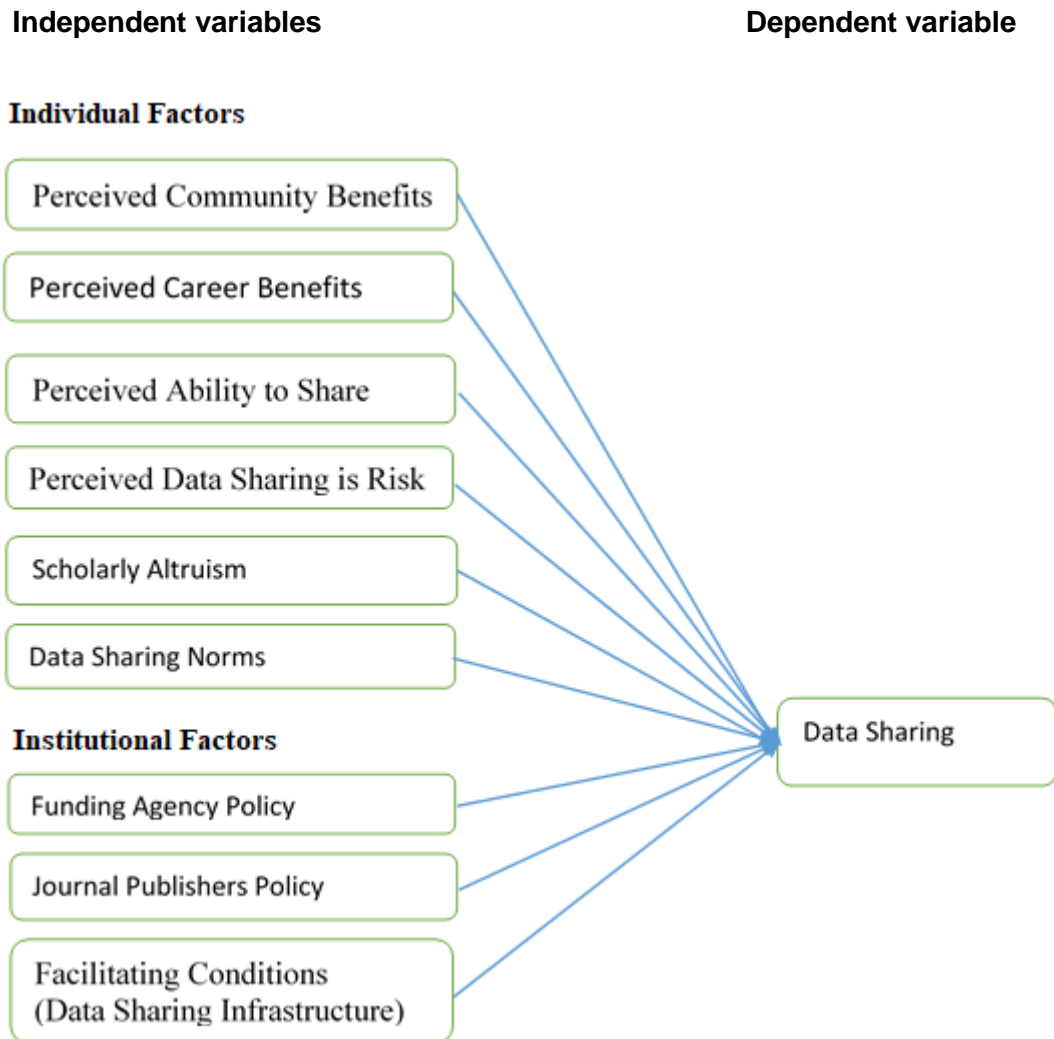


Figure 1: Research Data Sharing Conceptual Model

Source: Adapted and Modified from Ajzen (1991) and Ajzen & Fishbein (1980)

Based on the conceptual framework in Figure 2 below intervening variables comprise researchers and agricultural research institutions who receive and generate agricultural research data through research project activities. Researchers also use/reuse data in the research process, preserving and sharing their data. In sharing their data researchers are influenced by factors from the independent variable (institutional and individual factors) thus leading to actual (dependent variable) data sharing. Once data has been shared by depositing into data repositories, sending upon request and submitting as supplementary material in journal publication will enhance data accessibility. The shared data through different channels and other data-sharing infrastructures can be easily accessed by researchers for reused in future research.

Independent variables intervening variable Dependent variable

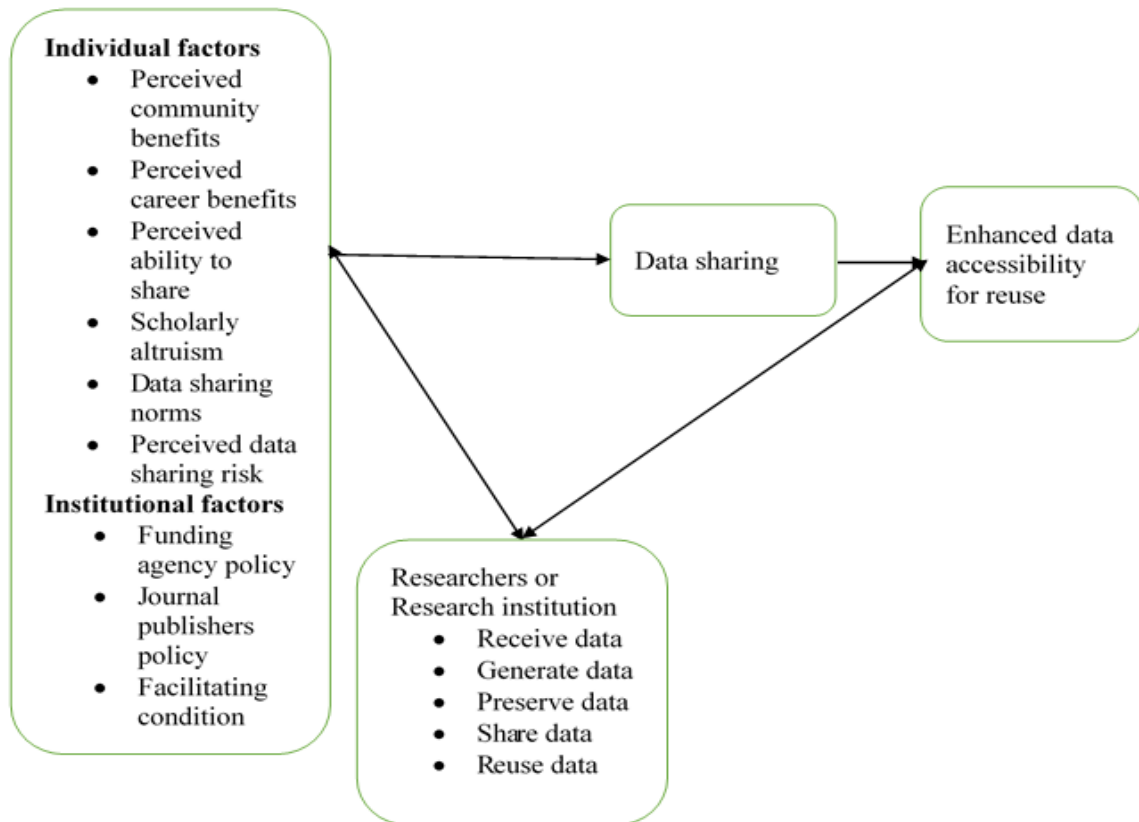


Figure 2: Conceptual framework showing data preservation, data sharing and reuse
Source: Designed by author 2023

1.6 General Methodology

This study is adapted to a pragmatic philosophy, which assumes that individuality may impact how people perceive the world, and therefore research is subjective. This theory uses both objective and subjective criteria (Al-Ababneh, 2020). The study employed a cross-sectional research design where the researcher surveyed and observed representative subjects at a specific point in time to examine their data-sharing practices and the factors that influence researchers in data sharing. The study applied both qualitative and quantitative approaches to data gathering and analysis. The qualitative approach was used to find out the views, insights and perceptions about data preservation, use/reuse, channels for data sharing and sharing practices among researchers. The quantitative approach was used to investigate the effectiveness of data preservation, data use/reuse, data sharing channels, sharing practices, and factors influencing researchers in data sharing.

1.6.1 Population and sampling

The target population of this research included agricultural researchers from different agricultural research fields working at the Tanzania Agriculture Research Institution (TARI) centres and at Sokoine University of Agriculture (SUA), specifically at the College of Agriculture and at the College of Veterinary Medicine, and Biomedical Sciences. The total population of researchers in the study areas was 527. From this study population, a

sample size of 227 was obtained using the formula for calculating sample size (Yamane, 1967).

To obtain the number of respondents to be involved in this study in each research institution a representative sample per institution was then calculated. The key informant interview was selected purposively, including 11 Directors/Managers from TARI centres and 1 Director of Postgraduate Studies, Research, Technology Transfer & Consultancy for SUA. Key-informants were purposively selected for interviews because they were thought to be more highly experienced and more knowledgeable and would therefore provide in-depth insights about the study topic.

1.6.2 Data collection

The data collection involved the use of multiple data-gathering techniques to examine the data-sharing practise and factors influencing researchers in their data-sharing. The cross-section survey method (questionnaire and interview) and focus group discussions (FGDs) were also employed in the study. The questionnaire in hard copies with both open- and close-ended questions was distributed and administered to 227 agricultural researchers from ten selected TARI centres, and SUA. Out of 227 distributed questionnaires 204 questionnaires were correctly filled and returned and used in this study.

Qualitative data were collected through interviews with key-informants and focus group discussions (FGD). Two focus group discussions with agricultural researchers were conducted at Mikocheni and Ilonga TARI Centres each focus group discussion had six participants selected randomly from among the survey respondents. These research institutions were selected randomly from among several agricultural research institutions. In addition, secondary data were collected from different sources, including the TARI communication strategy of 2019, the TARI Act of 2016, and the SUA Research Guideline and Regulation of 2019. The present study investigated the factors influencing agricultural researchers' data sharing practises for enhanced accessibility of data in Tanzania. To answer the research questions, a total of 204 valid responses were analysed. In the following section, the data analysis methods are presented.

1.6.3 Statistical data analysis

Quantitative data were analysed using descriptive statistical tools such as the statistical package for social science (SPSS) version 22. The quantitative analysis allowed researchers to obtain the frequency and percentages of the responses. The descriptive analysis allowed a researcher to describe data sharing, perceptions, and motivating factors for data sharing. While qualitative information was collected through interviews and focus groups were recorded, coded, and categorised into similar themes, explaining data-sharing practices among agricultural researchers, and were used in the discussion of findings. The inferential statistics was obtained by using the Binary Logistic Regression technique for testing the relationship between a dependent variable (data use/reuse) and independent variables (attitude, subjective norms, and facilitating conditions). Also, the technique was used in finding the relationship between individual and institutional factors that motivate researchers in data-sharing practices. This was performed to determine the survey variables that were statistically significant and those that were not statistically significant. The Binary Logistic Regression technique is used in this study because it measures the relationship between categorical variables against one or more independent variables. The dependent variables of this study (data use/reuse and data sharing) were categorical variable Yes and No whereas Yes =1 and No=0.

1.6.4 Measurement of construct

The proposed TPB theoretical research model was validated. Most of the variables were adopted from previous studies but a researcher added one variable perceived community benefits was modified to suit the context of this study. In this study, eight research constructs were measured. The constructs such as journal publisher's policies, scholarly altruism, perceived risk, and technical infrastructure were adopted from a previous study (Kim & Stanton, 2016). In addition, construct such as perceived career benefits, and perceived community benefits were adopted from another previous study (Yoon & Kim, 2019). Moreover, data-sharing norms and perceived behavioural control were also adopted from a previous study (Yoon & Kim, 2019; Zenk-Möltgen *et al.*, 2018).

In testing the existence or non-existence of a relationship between the dependent variable (data sharing) and independent variables (set of predictors), factors influencing data sharing, Logistic regression techniques were used the justification for this choice is due to the fact that the dependent variable was binary (Yes=0, No=1). Therefore, the statistical techniques were relevant to get the relationship between the independent variable and the dependent variable.

References

- Agrawal, A. (2017). *Data Roadmaps for Sustainable Development: Assessment and Lessons Learned (Issue December)*. Global Partnership for Sustainable Development Data, National Data Roadmaps Lesson Learned, 69pp. [http://www.data4sdgs.org/sites/default/files/services_files/RoadmapsAssessmentReport_Dec2017_final] site visited on 12/01/2022.
- Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behaviour and Human Decision Processes* 50(2): 179–211.
- Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behaviour*. Prentice-Hall. [<https://www.worldcat.org/title/understanding-attitudes-and-predicting-social-behaviour/oclc/5726878>] site visited on 29/05/2022.
- Al- Albahneh, M.M. (2020). Linking Ontology, Epistemology and Research Methodology. *Science and Philosophy* 8(1): 75-91.
- Aydinoglu, A. U., Dogan, G., & Taskin, Z. (2017). Research data management in Turkey: perceptions and practices. *Library Hi Tech* 35(2): 271–289.
- Bezuidenhout, L., & Chakauya, E. (2018). Hidden concerns of sharing research data by low / middle- income country scientists. *Global Bioethics* 29(1): 39–54.
- Bhatia, V., Stout, S., Baldwin, B., & Homer, D. (2016). *Results Data Initiatives: Findings from Tanzania*. Dar es Salaam: Development Gateway. [<https://org/developmentgateway.org/expertise/results>] site visited on 21/09/2021.
- Celli, F., Malapela, T., Wegner, K., Subirats, I., Kokoliou, E., Keizer, J., & Davey, R. P. (2022). AGRIS : providing access to agricultural research data exploiting open data on the web [version 1 ; peer review : 2 approved]. *F1000Research* 1–12.
- Chigwada, J., Chiparausha, B., & Kasiroori, J. (2017). Research data management in research institutions in Zimbabwe. *Data Science Journal* 16.
- COSTECH (2022). *National Research Registration and Clearance Guideline*, Tanzania Commission for Science and Technology (COSTECH), [<https://tawiri.or.tz/wp-content/uploads/2018/07/Research-registration-and-clearance-guideline-2018.pdf>] site visited on 27/10/2022.
- Curry, R. G., Crowston, K., Specht, A., Grant, B. W., & Dalton, E. D. (2017). Attitudes and norms affecting scientists ' data reuse. *PLoS ONE* 12(12):
- DPRTC (2019). *Sokoine University of Agriculture Research Regulations and Guidelines*, (4th ed.), (Issue June 2019), Morogoro: Tanzania. [<https://www.dprtc.sua.ac.tz/wp-content/uploads/SUA-Research-Regulations-and-Guidelines-2019.pdf>] site visited on 12/10/2021.
- Drakos, A., Protonotarios, V., & Manouselis, N. (2015). agINFRA : a research data hub for

agriculture , food and the environment [version 2 ; peer review : 2 approved].
F1000Research 4(127): 1–12.

Elsayed, A. M. (2018). Research Data Management and Sharing among Researchers in Arab Universities : An exploratory study. *International Journal of Bioassays* 44(4): 281–299.

FAO (2017). *Country Programming Framework for United Republic of Tanzania 2017-2020*, Prepared by The Government of the United Republic of Tanzania, Ministry of Agriculture, Livestock and Fisheries in collaboration with the Food and Agriculture Organization of the United Nations (FAO). [<https://www.fao.org/3/bt133e/bt133e.pdf>] site visited 26/11/2021.

Griffin, P. C., Khadake, J., Lemay, K. S., Lewis, S. E., Orchard, S., Pask, A., Pope, B., Roessner, U., Russell, K., Seemann, T., Treloar, A., Tyagi, S., Christiansen, J. H., Dayalan, S., Gladman, S., Hangartner, S. B., Hayden, H. L., Ho, W. W. H., Keeble-gagnère, G., & Schneider, M. V. (2022). Best practice data lifecycle approaches for the life sciences. [version 2 ; peer review : 2 approved]. site visited on 25/09/2022.

Gupta, S., & Muller-Bim, C. (2018). A study of e-Research and its relation with research data lifecycle: a literature perspective. *Benchmarking: An International Journal* 25(6): 1656–1680.

Ifeanyi-obi, C., & Ibisio, H. (2020). Creative Commons User License : CC BY-NC-ND. *Journal of Agricultural Extension* 24(4): 91–99.

Kamaruddin, I., Kadir, A., & Yunus, A. M. (2017). Recent Projects of a Preservation Framework for Digital Preservation. *International Journal of Academic Research in Business and Social Sciences* 7(11): 16–38.

Katabalwa, A. S., Bates, J., & Abbott, P. (2021). Potential opportunities and risks of sharing agricultural research data in Tanzania Keywords Research Environments in Tanzania. *IASSIST Quarterly* 45: 3–4.

Kijazi, A. (2018). *Data Availability, Exchange and Gaps: a case study of Tanzania*, [https://library.wmo.int/doc_num.php?explnum_id=5380] site visited on 25/4/2022.

Kim, Y., & Stanton, J. M. (2016). Institutional and Individual Factors Affecting Scientists Data-Sharing Behaviours: A Multilevel Analysis. *Journal of the Association for Information Science and Technology* 67(4): 776–799.

Kirub, A. (2016a). *Agricultural Research Data Management: Principles, Policy, and Practice*, Addis Ababa: Ethiopian Institute of Agriculture Research. [<http://197.156.72.153:8080/xmlui/bitstream/handle/123456789/217/data%20management-book.pdf?sequence=1&isallowed=y>] site visited on 03/02/2022.

Kirub, A. (2016b). *Agricultural Research Data Management: Principles, Policy, and Practice*, Addis Ababa: Ethiopian Institute of Agriculture Research.

[<http://197.156.72.153:8080/xmlui/bitstream/handle/123456789/217/data%20management-book.pdf?sequence=1&isallowed=y>] site visited on 03/02/2022.

- Kuteesa, A., & Kyotalimye, M. (2019). Documentation and data handling: How can Africa promote record keeping and investment in data management?. *African Journal of Food, Agriculture, Nutrition and Development* 19(1): 14171–14189.
- Late, E., & Kekalainen, J. (2020). Use and users of a social science research data archive. *PLoS ONE* 15(8): 1–17.
- Mboera, L.E. (2015). *The management of health and biomedical data in Tanzania, Need for a National Scientific Data Policy*, Directorate of Information Technology, National Institute for Medical Research, Dar es Salaam: Tanzania. [<https://slideplayer.com/slide/7625850/>] site visited on 21/04/2022.
- MUHAS (2020). *Data Sharing Policy*, Muhimbili University of Health and Allied Science (MUHAS). [<https://www.muhas.ac.tz/wp-content/uploads/2021/09/muhas-data-sharing-policy-2020.pdf>] site visited on 23/10/2022.
- Mushi, G. E., Pienaar, H., & Deventer, M. Van. (2020). *Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma, Tanzania*. *Data Science Journal* 19(1): 1-9.
- Mushi, G. E., van Deventer, M., & Pienaar, H. (2019). *Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma, Tanzania*. ResearchGate 1–19. <https://www.researchgate.net/publication/330853802> (site visited on 23/03/2021).
- Musker, R., & Schaap, B. (2022). Global Open Data in Agriculture and Nutrition (GODAN) initiative partner network analysis [version 1; peer review: 2 approved with reservations]. *F1000Research* 7(47).
- NBS (2010). Dissemination and Pricing Policy: The NBS recognizes that timely dissemination of data to specialized users for research purposes and other decision makers is beneficial. [<https://www.nbs.go.tz/index.php/en/about-us/policies-and-legislations/policies/12-Dissemination-and-pricing-policy>] site visited on 20/04/2022.
- Parr, C., Antognoli, E., & Sears, J. (2017). How Agricultural Researchers Share their Data: a Landscape Inventory. *Proceedings of TDWG 1*: e20434.
- Pronk, T. E. (2019). The Time Efficiency Gain in Sharing and Reuse of Research Data. *Data science journal* 18(10): 1-8.
- Schmidt, C., Hanne, J., Moore, J., Meesters, C., Ferrando-may, E., & Weidtkamp-peters, S. (2022). Research data management for bioimaging the 2021 nfdi4bioimage community survey [version 2; peer review: 2 approved]. *F1000Research* 11(638): 1-28.

- Sen, T. Z., Caccamo, M., Edwards, D., & Quesneville, H. (2022). Building a successful international research community through data sharing : The case of the Wheat Information System (WheatIS) [version 1 ; peer review : 2 approved , 1 approved with reservations]. *F1000Research* 9: 536.
- Shen, Y. (2017). Data sharing practices, information exchange behaviours , and knowledge discovery dynamics: a study of natural resources and environmental scientists. *Environmental Systems Research* 6(9): 1–14.
- Sielemann, K., Hafner, A., & Pucker, B. (2020). The reuse of public datasets in the life sciences: Potential risks and rewards. *PeerJ*8: 1–30.
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE* 6(6): 1–22.
- Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Danielle, P., & Kristina, D (2015). Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLoS ONE* 10(8): e0134826
- Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Research data sharing: Practices and attitudes of geophysicists. *Earth and Space Science* 5(12): 891–902.
- Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLoS ONE* 15(3): 1–26.
- TMA (2020). *Tanzania Meteorological Authority (TMA), Agro meteorological Database*. [http://agromet.meteo.go.tz/storage/docs/TMA_Agromet_Database_Manual.pdf] site visited on 12/02/2022.
- URT (2017a). *Information, Communication and Technology Policy (2017) for National Bureau of Statistics*. [<http://www.nbs.go.tz>] site visited 12/05/2022.
- URT (2017b). *The Journey to Better Data for Better Health in Tanzania: Tanzania Digital Health Investment Road Map 2017-2023*. [https://media.path.org/documents/DHS_health_tanzania_rpt1.pdfsite] visited 23/08/2022
- URT (2019). *Electronic Data Sharing and Exchange Guidelines. President's Office, Public Service Management and Good Governance*. [[https://www.ega.go.tz/uploads/standarddocuments/sw-1640950553-Final%20Electronic%20Data%20Sharing%20and%20Exchange%20Guidelines_EDITTED%20\(1\)](https://www.ega.go.tz/uploads/standarddocuments/sw-1640950553-Final%20Electronic%20Data%20Sharing%20and%20Exchange%20Guidelines_EDITTED%20(1).)].
- URT (2016). *Tanzania Agricultural Research Institute (TARI) ACT No 10 of 2016*, Dodoma: Tanzania. [https://www.tanzania.go.tz/egov_uploads/documents/sheria_chapa_-Tanzania_Agricultural_Research_inst_sw.pdf] (site visited on 13/2/2022).

- Van Bell, J., Iglesias, C., Mungal, P., Lammerhirt, D., Nuhu, H., Hlabano, M., Nest-Nash, T., & Chaudh, S. (2018). *Africa Data Revolution Report 2018*. <http://webfoundation.org/docs/2018/10/African-data-revolution-report-brochure-publish.pdf> (site visited 17/10/2022).
- Williams, S. C., Farrell, S. L., Kerby, E. E., & Kocher, M. (2019). Agricultural Researchers' Attitudes Toward Open Access and Data Sharing. *Issues in Science and Technology Librarianship (ISTL)* 91: 1–16. [<https://journals.library.ualberta.ca/istl/index.php/istl/article/view/4/47>] site visited on 12/07/2022.
- Williamson, H. F., Brettschneider, J., Caccamo, M., Davey, R. P., Goble, C., Kersey, P. J., May, S., Morris, R. J., Ostler, R., Pridmore, T., Rawlings, C., Studholme, D., Tsafaris, S. A., & Leonelli, S. (2021). Data management challenges for artificial intelligence in plant and agricultural research [version 1 ; peer review : 1 approved with reservations]. *F1000Research* 10(324): 1–20.
- Yamane, T. (1967). *Statistics: An Introductory Analysis* (2nd. ed.), New York: Harper and Row.
- Yoon, A., & Kim, Y. (2019). The role of data-reuse experience in biological scientists' data sharing : an empirical analysis. *The Electronic Library* 38(1): 186–208.
- Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V., & Balaban, E. (2018). Factors influencing the data sharing behaviour of researchers in sociology and political science. *Journal of Documentation* 74(5): 1053–1073.
- Zoran, S. (2017) Research and data-sharing policy in Sweden– neoliberal courses, forces and discourses. *Prometheus* 35 (4): 249-266.
- Zhao, H., & Wang, J. (2015). Development and current situation of agricultural scientific data sharing in China. Agricultural Information Institute of CAAS, Beijing. *IFIP Advances in Information and Communication Technology* 452: 80–86.

CHAPTER TWO

PAPER I

[check for updates](#)

Article



Data preservation practices for enhancing agricultural research data usage among agricultural researchers in Tanzania

Journal of Librarianship and Information Science

1–13

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09610006221138110

journals.sagepub.com/home/lis


Nolasko Victory Mwinami

Sokoine University of Agriculture, United Republic of Tanzania

Frankwell W. Dulle

Sokoine University of Agriculture, United Republic of Tanzania

Wulystan Pius Mtega

Sokoine University of Agriculture, United Republic of Tanzania

Abstract

The objective of this study was to investigate the role of research data preservation for enhanced data usage among agricultural researchers in Tanzania. Specifically, the study aimed to examine the data preservation methods used by agriculture researchers, find out how long agriculture researchers preserve their agriculture research data, and determine factors that influence agriculture researchers on their choice of data preservation methods for use. The study employed a cross-sectional research design. The study employed both qualitative and quantitative approaches. A survey was conducted to collect data in 11 research institutions. A simple random sampling technique was used to select 204 respondents from the study area while purposive sampling techniques were used to select 11 agriculture research institutions including 10 Tanzanian Agricultural Research Institution (TARI) centers, and Sokoine University of Agriculture (SUA). Also, 12 respondents were selected purposively for an in-depth interview as key informants. The study adopted Data Curation Centre (DCC) Lifecycle Model to explain data preservation process. Findings indicated that a majority of more than 90% of researchers preferred to preserve their data using different storage devices such as field notebooks, computers, and institutional libraries. Moreover, findings indicated that about 74% of agricultural researchers preferred to preserve their data for more than 6 years after the end of the project. Findings also indicated factors that influence researchers in the choice of data preservation methods to be easy to reach, cost-effective storage devices, support to use the devices, adequate infrastructure for data preservation, and reliable power supply. It can be concluded that there is yet a great role of research data preservation in enhancing data usage among researchers in Tanzania. It is recommended that the government should establish an agricultural research data bank to guarantee permanent availability of data at all times when needed.

Keywords

Agricultural research data, data preservation, data reuse, data sharing, data storage

Introduction

Preservation of agricultural research data by research institutions is vital for enhanced sharing, reuse, and development of the agriculture sector. However, this depends on the choice of preservation method used. According to Kirub (2016), it is the primary role of Agriculture research institutes to collect, generate, store, and use these data. However, such data are faced with a threat of being lost at

the end of the project. Globally, it has been advocated that proper preservation of such data can guarantee easy access, browsing, consultation, and future use (Tripathi et al., 2017), and also can facilitate technology transfer, and

Corresponding author:

Nolasko Victory Mwinami, Sokoine University of Agriculture, P. O. Box 3158, Morogoro, 67112, Tanzania.
Email: mwinaminolasko@yahoo.com

innovation in agriculture (Ng'eno and Mutula, 2018a), creation of opportunities for access, sharing, and reuse by other researchers beyond a given geographical area (Brouder et al., 2019; Dileepkumar, 2014; Eckes et al., 2017; Granda and Blasczyk, 2011; Kirub, 2016; Zhao and Wang, 2015). The choice of the best preservation method is a pre-requisites for its success (Ng'eno and Mutula, 2018a). Data preservation practice has attracted more attention from research institutes, researchers, funding institutions, and international development agencies (Adrian et al., 2018; Kirub, 2016).

Some funding agencies and publisher advocate the deposition of data on open access repositories to facilitate easy access, and use/reuse by others (Tripathi et al., 2017). Likewise, the government of Tanzania has taken several initiatives to preserve different types of data. The establishment of electronic platforms at the National Bureau of Statistics (NBS), and Tanzania Meteorological Agency (TMA) facilitated the preservation and dissemination of scientific, baseline survey, and agriculture administrative data (Agrawal, 2017; Bhatia et al., 2016; TMA, 2020; URT, 2017). While these initiatives have been successful, there are limited initiatives that have focused on preserving agricultural research data in agriculture research institutions in Tanzania. Moreover, the SUA research regulation and guidelines of 2019 and TARI Act of 2016 both emphasize the preservation of research data and other research materials but the documents are silent on methods of data preservation for ensuring long-term data access, sharing, and future reuse (DPRTC, 2019; URT, 2016). Despite previous literature available on data preservation in the study area (Agrawal, 2017; Chidi, 2017; FAO, 2017; Katabalwa et al., 2021; Kijazi, 2018; Mboera, 2015; Mushi et al., 2020; NBS, 2010; TMA, 2020), however, there is a paucity of information on agricultural research data preservation in agricultural research institutions in Tanzania. This study, therefore, aims at closing this information gap.

Objectives of this study

The primary objective of this study was to investigate the roles of agriculture research data preservation methods/devices for enhancing data usage among researchers in Tanzania.

The specific objectives of this study

1. Examine the methods for the preservation of research data from research institutions in Tanzania.
2. Find out for how long agriculture researchers preserve their agriculture research data.
3. Determine factors influencing researchers in their choices of data preservation methods.

Significance of the study

The study will enlighten the agricultural research institutions to strengthen existing and adopt better methods for preserving the agricultural research data in short-term and long terms plans. Findings will create awareness in agriculture research institutions to see the need of strengthening the department of technology transfer to perform a coordination function to manage and preserve agricultural research data. The study shall also contribute to the existing body of knowledge by identifying data preservation methods, and factors that influence researchers on the choice of data preservation methods.

Literature review

This section organized literature related to the study topic. The literature review is organized around broad themes based on study objectives as follows: scientific data preservation methods, duration of data preservation, and factors that influence agriculture researchers in their choices of storage and preservation methods.

Research data preservation

Data can be in form of images, recordings, musical compositions, verbal communications, experimental readings, simulations, codes, textual, qualitative, and quantitative (Tripathi et al., 2017). While research data refers to the basic data which are generated by a human being through technological activities for example structure graphic data (SPSS, GIS), source code data, images data (JPEG, GIF, TIFF) (Adika and Kwanya, 2020; Zhao and Wang, 2015). In this case, agricultural research data can exist in the form of text documents, graphics, databases, spreadsheets, laboratory notes, films, photographs, test responses, and audio (Chigwada et al., 2017; Elsayed and Saleh, 2018). Furthermore, agricultural researchers generate several types of data such as crop breeding data, baseline survey data, genomic data, remote sensing data, and geographic information system (GIS) data (Dileepkumar, 2014; Kirub, 2016). Such important data generated are required to be preserved for future reuse by researchers and other stakeholders.

Data storage is among the activities in the data management life cycle. Adika and Kwanya (2020) define Research Data Management as a set of activities that include creation, organization, structuring, naming, baking up, storage, conservation, and sharing of data. Therefore, data storage is a process of recording or locating data in a medium that is managed by an individual or institution that allows access and sharing (Frederick et al., 2019). Such medium may include all storage devices, storage on a network, or systems, and storage in the cloud such as Amazon Web Services and Google Cloud Platform. The term data

preservation is defined as the permanent protection of research data that has been obtained from the finished project to guarantee data access for future use (Kruse and Thestrup, 2014) as cited in (Ng'eno and Mutula, 2018b). In the same vein, Mushi et al. (2020) pointed out that data preservation is one of the important activities carried out in the Research Data Management (RDM) life cycle though it has not yet been implemented in African countries including Tanzania.

Previous studies have shown that data can be preserved for the short term or long term. Literature shows that researchers prefer to store their data in short term during the project life cycle while long-term data storage was done after the end of the project (Dileepkumar, 2014; Tenopir et al., 2020). Researchers prefer to store their data temporally in their personal computers during the project life cycle. In line with this, Tenopir et al. (2020) exposed that researchers from Space and Planetary Science were satisfied to store data in their personal computers, departmental servers, or USB driver during the project life cycle. In the same study other respondents belonging to Marine Ocean (55.8%), Biology (50.9%), and Agriculture and Natural Resources (52.1%) were satisfied with long-term data preservation. Therefore, researchers are storing their data in short, and long-term data storage. The study also examined research data management in Kenya's agricultural research institutes (Ng'eno and Mutula, 2018b) findings revealed that the majority of respondents (20.2%) kept their research data for a period of more than 10 years, and those who kept their data between 5 and 10 years were 20.2%. Other respondents (23.4%) from this study did not know the length of time for research data preservation in agricultural research institutes.

Research data preservation methods

Previous literature has shown that researchers use different methods to preserve their data. Agricultural researchers from National Agriculture Research Systems (NARS) in Sub-Saharan Africa and lecturers at Strathmore University in Nairobi, Kenya were reported to use personal computers, flash, and external hard disks in managing their research data (Adika and Kwanya, 2020; Maru, 2004; Sielemann et al., 2020). Moreover, in a study that examine research data management in Kenya's Agricultural Research Institute (Ng'eno and Mutula, 2018a) it was found that (86.3%) preferred to preserve their research data on personal computers, and about 79% preferred to store their data into the hard drive, and 94.5% reported to preserve their data into their departmental servers, and institutional repository.

In line with this, in the study that examined data sharing, management, use, and reuse practices and perceptions of scientists worldwide (Tenopir et al., 2020) findings

exposed that 61.3% of respondents stored their data on personal computers, followed by 42.9% stored in institutional servers and 29.8% of respondents stored their data using their USB/external drivers. Likewise, in the web-based survey that investigated data preservation at the University of Kansas (Weller and Monroe-Gulick, 2014) findings showed that researchers preserved their data using electronic devices, and tools such as hard drives, CDs, and cloud-based storage like Dropbox or Google Drive and centralized server. This implies that the researcher had more options for data storage and preservation. However, these findings show that data stored on a flash drive and a personal computer was a bad practice while a good practice was storing data in institutional repositories and departmental servers. Data storage through personal devices was viewed as bad practice simply because it denies data access to others.

Literature has revealed that researchers prefer to preserve their data on a centralized data platform. The ICRISAT program which operates in Asia and Sub-Saharan Africa came up with innovative data preservation, publication, and web-based sharing platforms. Through this program, several data preservation platforms were initiated including Dataverse, Cloud-based M, AGROBESE, and Intergraded Breeding Platform (IBP) (Dileepkumar, 2014; Maru, 2004). These data platforms provided a chance for agricultural researchers elsewhere to access, share and reuse other researchers' data. In line with this, the European Biodiversity Network (EU BON) adopted a tool for data preservation, sharing, and description. In this network, several electronic data management platforms were established including Drupal Ecological Information Management System, Metacat and Morpho, Plazi Treatment Bank, Golden Gate Imagine, Plouto, and GBIF Intergraded Publishing Toolkits. These data preservation platforms enabled researchers to publish, share and access biodiversity data for use (Smirnova et al., 2016). These data platforms facilitated data sharing and easy data accessibility among members.

Moreover, research guideline from SUA and TARI headquarters stipulates that research data and other materials generated by researchers and students belong to the research institutions, and must be submitted to be preserved in the research institutions (DPRTC, 2019; URT, 2016). However, research guidelines do not tell something about how such important data are preserved for future usage. None of these previous studies managed to give a clear comprehensive understanding of research data preservation methods used in agricultural research institutions in Tanzania.

Choices for data preservation method

Several factors can influence researchers in the selection of data preservation methods.

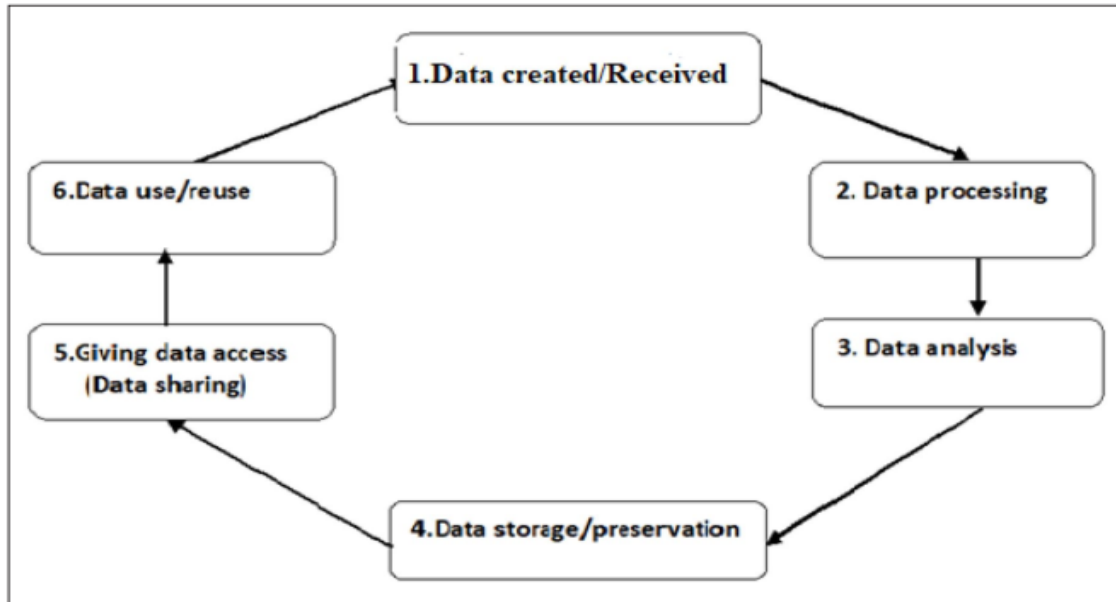


Figure 1. Agricultural research data life cycle model.
Source: Adapted and modified from Higgins (2008).

Previous literature has revealed that the availability of Information Communication Technology (ICT) tools and infrastructures in agricultural research institutions facilitated researchers to capture data, store, manipulate, analyze, preserve, and share their research data (Ng'eno and Mutula, 2018a; Zhao and Wang, 2015). Furthermore, technical support provided to researchers enhances researchers to choose the appropriate method for data preservation. Tenopir et al. (2020) revealed that the availability of data managers or data librarians who were readily available to assist researchers in their data preservations could facilitate researchers to preserve their data in long term by depositing their data into repositories beyond the project life cycle.

Furthermore, the cost of data preservation and management technology can encourage or discourage researchers to use a particular technology. If the cost of using the device is cheaper can influence researchers to use a certain dataset also if the cost is higher researchers will be discouraged to use the device or method to manage their data (Kirub, 2016; Musker and Schaap, 2018). Likewise, data preservation devices use electricity. The availability of reliable electricity is one of the factors influencing the choice of method to preserve data. Electricity is important in running electronic devices, and equipment. Unreliable electricity supply may hinder researchers not being able to use a particular data preservation equipment. According Barakabitze et al. (2015) reported that there were unreliable electric supplies to operate computers and other electronic storage devices in agricultural research institutions in Tanzania.

These previous studies have indicated the existence of some factors that influence researchers on the choice of method or device to use in preserving data. The methods varied depending on the duration of data preservation. Researchers preferred to preserve their data using personal preservation devices for the short term. While in some institutions researchers preferred to use formal or centralized systems to preserve their data for the long term beyond the project lifecycle. It is not known yet how agricultural researchers in Tanzania preserve their data during and beyond their project life cycle. Due to this scant documented evidence, there was a need to fill this knowledge gap.

Theoretical framework

This study adopted Data Curation Centre (DCC) Life cycle Model Figure 1 below. The Curation Lifecycle Model was proposed by the Digital Curation Centre (DCC) to promote a life cycle approach to the management of digital materials, to enable their successful preservation from the initial stage to their disposal or selection for reuse and long-term preservation (Higgins, 2008). The DCC model was useful to guide this study because the model explains that data are created, received, analyzed, preserved, shared, and finally data act as raw materials to generate new data Figure 1 below. The model is relevant to this study because it clarifies data preservation as one of the core activities in the data life cycle. Data preservation or management is a participatory activity it involve researchers, data managers, curators,

repository specialist, librarians, and project managers (Pouchard, 2016). The model expresses the value of data preservation thus without data preservation, data will not be available, easily accessed for used/reused in future research (Ng'eno and Mutula, 2018).

The Data Curation Centre (DCC) Lifecycle Model in Figure 1 below includes the data capture, processing, analysis, preservation, access, re-use, and transformation of research data (Ng'eno and Mutula, 2018a).

Data creation/capture

Data are created through the research process. Also, data may be captured by receiving data from other sources (Adika and Kwanya, 2020).

Data analysis

After data is created/received then data analysis is carried out. After analysis data can be preserved for future use and reuse while other data can be published in scientific works.

Data preservation

Data preservation means protecting data in a secured environment for the long term that will allow future access and reuse (Ng'eno and Mutula, 2018b). Researchers should choose better devices or methods for better data preservation. Better choice of devices or methods can allow data access and reuse in the future.

Data access

Data may be accessed from storage devices or methods. Data will be accessed when they are shared into accessible sources such as personal storage devices (computers), databases (data repositories), and libraries.

Data use/reuse

The preserved data may be used and reused if they were preserved on the devices or a method that allows data access and usage (Higgins, 2008).

Agricultural research data preservation conceptual framework

The conceptual framework in Figure 2 below indicates researchers create data during research activity or project (Tripathi et al., 2017). Also, researchers receives data from other sources to be used in the research process. Data obtained from research are analyzed, interpreted, and shared through publications while other data are preserved for future use/reuse. Researchers are required to have

better choices in the selection of data preservation devices or methods. The right choice of data storage and preservation devices/methods will ensure the data is effectively preserved and used. The data that has been preserved using a better storage device/method will enhance data access, sharing, use, and future reuse (Hawkins et al., 2018; Higgins, 2008).

Methodology

The study employed a descriptive and cross-sectional research design to investigate the role of research data storage and preservation for enhanced data usage among agricultural researchers in Tanzania. The survey was conducted from February 2020 to March 2022. The study used both qualitative and quantitative approaches in data collection and analysis. The qualitative approach was used to determine opinions and views on data storage and preservation. While the quantitative approach was used to examine the preservation methods and factors influencing researchers on the choice of storage and preservation methods/devices.

A combination of the two approaches helped a researcher to get relevant data for the study. The study was conducted in six agricultural research institutions in Tanzania. The ecological zones included in the study were randomly selected. These included; Eastern, Central, Northern, Southern, Southern Highlands, and Lake Zone. Tanzania Agriculture Research Institution (TARI) centers and SUA specifically the College of Agriculture and College of Veterinary Medicine and Biomedical Sciences were purposively selected to be included in the study. The selection of these research institutions considered the agricultural research institutions which had more than twenty (20) researchers at the institution. This number of researchers was enough to justify the data collection process.

The study included agricultural researchers. The total population of researchers in the study areas was 527. From this study population, the sample size of 227 was obtained using Yamane's (1967) formula for sample size calculation.

$$n = \frac{N}{1 + N(e)^2}$$

Where:

n= Sample size (227)

N= Total study population (527)

e= Level of significance (5%)

After obtaining the sample size of 227 then a representative sample per institution was calculated. This representative sample size per institution facilitated the determination of the total number of respondents to distribute the questionnaires. The representative sample per institution was

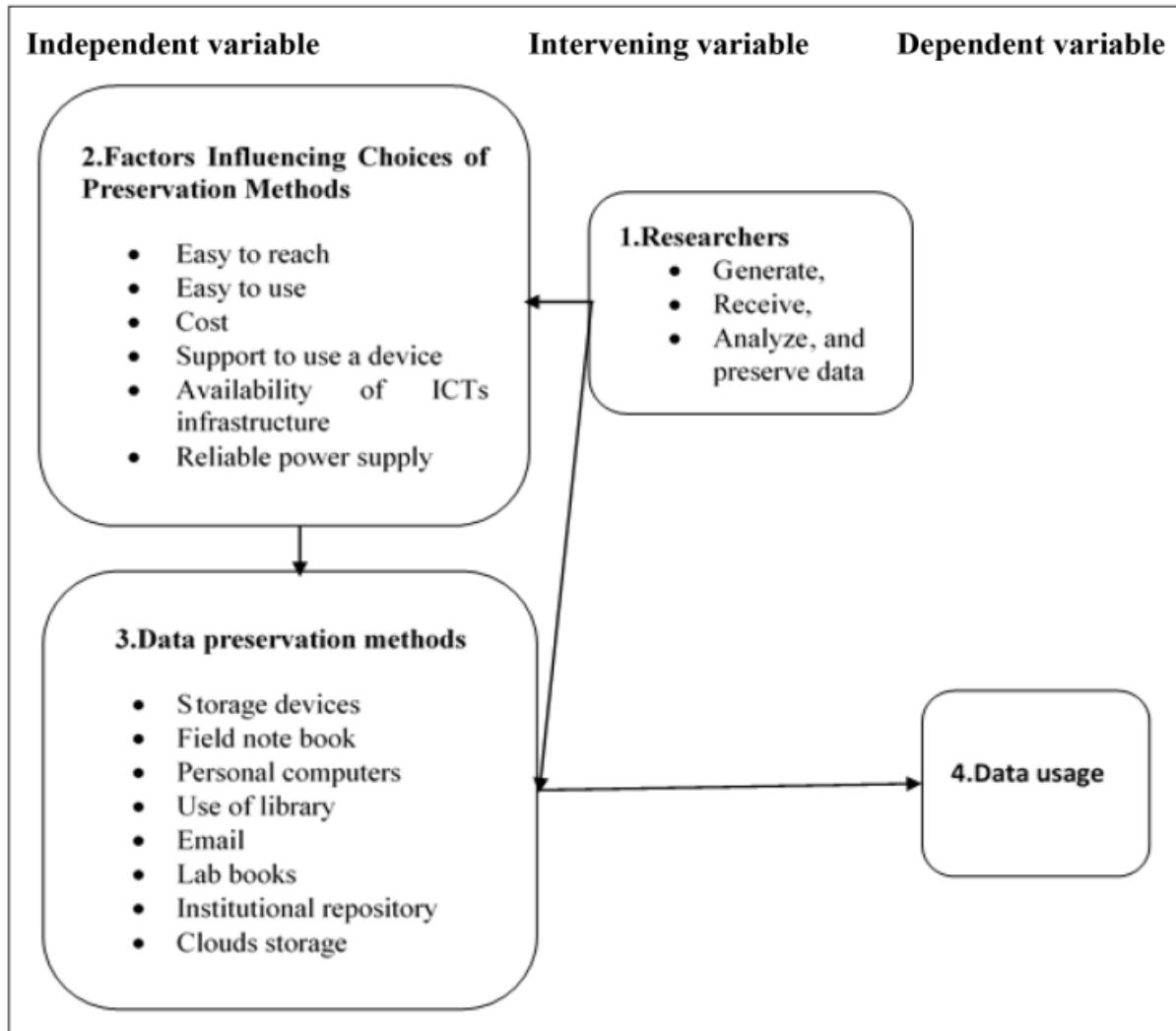


Figure 2. Agricultural research data preservation conceptual framework.
Source: Designed by researcher, 2022.

calculated (Table 1). From the sample size per institution, the respondents were randomly chosen to represent the 227 sample size. Eleven directors from TARI and one from SUA Post Graduate Studies were selected for key informant interviews. These were purposively selected because they were thought to be experienced and knowledgeable and would therefore provide in-depth insights about the topic.

Data collection involved the use of multiple data gathering techniques to investigate the methods use in data preservation among researchers. The cross-section survey method (questionnaire and interview) and focus group discussions were employed in the study. The questionnaire with both open and closed-ended questions was distributed to 227 agricultural researchers from SUA and 10 selected TARI centers. A total number of 204 questionnaires from

SUA and 10 selected TARI centers were correctly filled, returned, and used in this study (Table 1). The qualitative data were collected through interviews with key informants that lasted 15–20 minutes.

A total of two focus group discussions with agricultural researchers were conducted at Ilonga and Mikocheni centers, and each group had five participants. These research institutions were selected randomly from among several agricultural research institutions. The main purpose was to understand the types of research data preservation that are used in sharing agriculture research data. During the discussion respondents were free to express their views which were recorded by the researcher. In addition, secondary data was collected from different sources, including the SUA Research Guideline and Regulation of 2019 and the TARI Act of 2016.

Table 1. Sampling frame ($n = 204$).

Zone	Institution	Total population	Sample size per institution (S) ^a	Returned questionnaire
Eastern	SUA	223	96	84
Lake	Ukiriguru	42	18	15
Southern highlands	Uyole	41	18	17
Southern	Naliendele	38	16	15
Northern	Selian	37	16	16
Eastern	Mikocheni	30	13	11
Eastern	Ilonga	28	12	12
Central	Makutopora	26	11	10
Eastern	Kibaha	22	9	9
Eastern	Katrin	20	9	8
Northern	Tengeru	20	9	8
	Total	527	227	204

Source: Field data, 2021, sample size per institution.

^aStudy sample size (SS) (227) times population per institution over total population (TT).

Pre-test

Before embarking on data collection, the instruments were pre-tested by 30 respondents in the College of Forestry, Wildlife, and Nature Conservation at SUA. The copies of the questionnaire were sent to them to check the overall structure of wording, presentation, and the relevance of the questions concerning the study objectives. Based on Cronbach's Alpha for testing the reliability of research tools, the test result gave an alpha of 0.943 which was above the 0.70 acceptable standards meaning that the tool was reliable to be used. A pre-test helps to identify errors that a researcher may have not foreseen during designing a collection tool (Plooy-Cilliers et al., 2014). Concerning ethical considerations, the researcher avoided any risk of harm during data collection and ensured the right of privacy and informed consent before engaging any respondent in the study.

Data collection and analysis

The data that was collected and analyzed included both qualitative and quantitative data. Qualitative analysis was accomplished using content analysis, the content analysis was used to analyze qualitative data that was obtained from key informant interviews, and Focus Group Discussions (FGDs). Quantitative analysis was accomplished by using statistical analysis, which is a technique that enables a researcher to examine quantitative data with the help of computer software packages for data analysis (Methew and Ross, 2010). The Statistical Product and Service Solution (SPSS) software version 22 and Excel software were used to carry out a descriptive analysis. The descriptive statistics results such as frequencies and percentages facilitated to organize and summarize data on methods used in data preservation, length of data preservation, and factors influencing researchers on choices of data preservation method.

Results and discussion

This part presents social-demographic characteristics of respondents such as sex, age, education level, research experience, and the research area of specialization. Also, this part presents findings based on study objectives; data preservation methods used to enhance sharing, and the factors influencing the researcher's choice of preservation methods.

Characteristics of respondents

The social-economic characteristics of the respondents are given in Table 2. Results show that there were twice as many males (67%) and (33%) female which shows that males dominated in agricultural research activities than females. Similarly, findings showed that 37% of respondents were in the age range of 31–40 years. Findings imply that there were more male agricultural researchers than female researchers within the institutions under study. This situation has been a tradition in many agricultural research institutes and in universities that have specialization in agriculture. It is not surprising to find more male employees than female employees. It is also worth noting that all of these institutions are dealing with scientific inventions and discoveries whose implementations require a background in science subjects. The findings are related to the study that found there were fewer women than men trained, recruited, and employed in the agricultural sciences and jobs related to science, technology, engineering, and mathematics (STEM) (Funk et al., 2018).

In terms of age distribution majority (37%) of respondents had an age distribution ranging from 31 to 40 followed by (32%) of respondents who had an age distribution ranging from 41 to 50 years. Respondents about (27%) with ages ranging from 51 to 61. This study found that the majority of the researchers in agricultural research institutions were aged 31 years and above. These results imply

Table 2. Demographic characteristics of respondents (n = 204).

Gender	Frequency	Percentage
Male	137	67
Female	67	33
Age		
More than 61 years	2	1
51–60 years	55	27
41–50 years	66	32
31–40 years	75	37
21–30 years	6	3
Education level		
PhD	45	22
Master degree	123	60
Bachelor degree	36	18
Work experiences		
More than 8 years	163	79.9
6–8 years	17	8.3
3–5 years	20	9.8
Lesser than 3 years	4	2.0

Source: Field data, 2021.

that researchers from 31 years and above have been working on research projects for substantial years and therefore have generated many agricultural research data that are needed to be preserved for future use.

Findings in Table 2 indicate that majority (60%) of respondents had Master's degree while 22% of respondents had a PhD and 18% of respondents had a Bachelor's degree. In terms of education level of respondents, findings indicate that agricultural researchers are highly qualified for their work as the majority had a Master's degree and above in agricultural sciences. The respondents with this academic qualification have enough knowledge and skills in conducting research and preserving their research data.

Furthermore, findings in Table 2 below indicate that 87% of respondents were those with work experiences of more than eight (8) years. While about 11% of the respondents were those with research experiences between 3 and 5 years. About 8.3% of respondents were those with research experience between 6 and 8 years. And finally, 2% of respondents were those with lesser than 3 years of experience. Concerning work experience in research activities, findings signify that agricultural research activities require researchers who are well experienced to achieve the intended goal of improving agricultural production. Many agricultural researchers from SUA and TARI centers are well-experienced researchers they have a lot of research data that are preserved in their devices for future use and reuse. This demographic information was useful to indicate the suitability of the units for the study.

Findings in Table 3 below indicate the researchers' academic field of specialization. From the findings, the

Table 3. Researchers academic field of specialization (n = 204).

Field of specialization	Frequency	Percentage
Botany, crop sciences, and horticulture	92	45
Animal science, aquaculture, and range management	16	7.8
Soil science and geological sciences	12	5.9
Veterinary medicine, surgery, and public health	11	5.4
Agriculture economics	11	5.4
Food technology, nutrition, and consumer sciences	10	4.9
Engineering sciences and technology	9	4.4
Agriculture biotechnology	9	4.4
Environmental science	8	3.9
Agro ecology	3	1.5
Public health and food safety	3	1.5
Natural resources management	2	1.0
Epistemology and laboratory management	1	0.5
Hydrology and water resources engineering	1	0.5
Total	204	100.0

Source: Field data, 2021.

majority (45%) of respondents indicated specializing in botany, crop sciences, and horticulture. In addition (7.8%) of respondents specialized in Animal Science, Aquaculture and Range Management, and Agricultural Extension and Community Development. Other respondents specialized in various fields as shown in Table 3 below. These results indicate that researchers from all research institutions specialized in one or more research fields. The primary role of research institutions has been stipulated in the Act of 2016 for TARI part 4(1) and from SUA Research Regulations and Guideline (DPRTC, 2019; URT, 2016).

Research data storage and preservation methods

The Data Curation Centre (DCC) Life cycle Model as shown in Figure 1 above emphasizes data storage and preservation as among the core activities in the DCC life cycle. Data preservation involves several data preservation methods. The finding in Table 4 below shows that more than 80% of respondents indicated they preserved their data using several methods/devices such as (flash sticks, external hard disks, CDs, and DVDs), use of field notebooks, and desktop computers found in departmental offices, and personal laptops. The findings in Table 4 below indicate that the prominent methods used by agricultural researchers to preserve their data include the use of storage devices such as (flash sticks, external hard disks, CDs, and DVDs), use of field notes books, and computer desktops and laptops. These findings imply that

Table 4. Data preservation methods existing and used in research institutions ($n = 204$).

Storage method	Frequency	Percentage
Storage device (flash sticks, external hard disk, CDs, and DVDs)	159	77.9
Field notes books	171	83.8
Computers desk/laptops	167	81.9
Institutional library	160	78.4
Storage device (flash sticks, external hard disk, CDs, and DVDs)	159	77.9
Use of email	107	52.5
Use lab books	91	44.6
Institutional database/repository	59	28.9
Use of clouds storage	50	24.5

Source: Field data, 2021.

researchers were preserving their data using informal ways that do not guarantee data availability and accessibility by all researchers. Findings have revealed that researchers from different research institutions preserve their data in their ways, and it happens the absence of the staff for any reasons may hinder data access among researchers. This practice is not health for the progress of science because researchers may lack important data for future research. However, some data preservation methods were not prominently used by agricultural researchers including lab books, institutional databases/repositories, and cloud storage. Among the reasons for the rare usage of these data preservation methods included low literacy in the use of data preservation methods such as cloud storage and the absence of data storage infrastructure such as institutional data repositories in the study area. SUA had an institutional repository that is not specified as a data repository however it hosts several types of research data in the form of research outputs including thesis/dissertation, and other electronic resources.

In addition through FGD from the Ilonga TARI center respondents reported:

We preserve our research data using our computer laptops/desktops, flash sticks, external hard disks, and CDs/DVDs. We submit our technical report to the Tanzania Agriculture Research Institute (TARI) General Director's Office for permanent preservation. Also, we normally submit research data to the project funder through their platform that is opened specifically for that project. At the end of the project sometimes you cannot be able to log in to that electronic platform. In our institute, we do not have a centralized database or data repository for research data preservation.

In addition through the interview from the TARI Ukiriguru center respondents reported:

I preserve my data in the departmental desktop computers as a common poor for permanent data storage. Also, at the end

of the research projects, I submit research data and other substantial research reports in hardcopies, and softcopies to the technology transfer department for permanent preservation.

In addition through an interview from SUA respondents reported:

Until now we do not have a data repository though we have an institutional repository used for preserving research outputs such as theses submitted to the institutions by researchers (academic staff), and students.

The use of personal laptops, other personal storage devices, and field notebook to preserve data is not a good practice this hinder data accessibility by other researchers. Moreover, it decentralizes data management to persons not trained as data storage managers (none data storage professionals). These findings are related to the previous studies which indicated that (22.6%), (8.8%), and (17%) of respondents for research institute A, B, and C respectively preferred to store their data in short term through their laptops, CDs, USB drives (Ng'eno and Mutula, 2018a; Tenopir et al., 2020). Furthermore, Aydinoglu et al. (2017) revealed that academics in Turkey used several ways to back up their data including CD/DVD/external hard disk and thumb drive, and cloud storage. There are some disadvantages of data storage on the hard disk and other magnetic devices. Data storage in magnetic tape and hard disk storage medium does not provide guarantees over time because it would become demagnetized and corrupted (Frederick et al., 2019). Therefore, the absence of research data repositories in agricultural research institutions in Tanzania might hinder researchers to have easy access, sharing, use, and reuse of their data.

Length of data preservation in research institutions

Table 5 below indicate the length of data preservation in agriculture research institution during, and beyond the project life cycle. The majority (74%) of respondents indicated they preserve their research data for more than 6 years. These findings imply that researchers preserved data for more than 6 years to enhance effective data use and future reuse. For example, researchers could prefer to use a computer, flash memory, and CDs to store data temporarily during the project life cycle while after the end of the project researchers could preserve their data in the office computers and also publish their data in funders database, online journals, submit research findings, and research materials to the institutions. Hawkins et al. (2018) revealed that cloud storage of agriculture data is a good way to store data for long-term stewardship.

Table 5. Length of research data preservation in research institutions ($n = 204$).

Responses	Frequency	Percent
Lesser than 3 years	4	2.0
1–3 years	16	7.8
4–6 years	33	16.2
More than 6 years	151	74.0
Total	204	100

Source: Field data, 2021.

Factors influencing researcher's choice of preservation methods

Table 6 below show results from the descriptive data analysis showing the factors that influence the choice of data preservation method/device. Findings as shown in Table 6 below, indicated that the majority (50.3%) of respondents indicated agreed and (38.7%) indicated strongly agree that easy-to-reach storage method/device influenced agricultural researchers choice of method. These findings imply that researchers preferred to use storage devices, field notebooks, and personal computer laptops because they were easy to reach. It can be said that almost every researcher possesses a computer laptop and other storage devices so it is easy to reach and use such devices. These findings are related to a prior study (Weller and Monroe-Gulick, 2014) revealed that researchers preserved their data through hard drives, CDs, and university servers because they were easy to reach and use. Similarly, Tenopir et al. (2020) found that researchers were more satisfied with the availability of data in the repository because it was easy for them to locate and access data.

Furthermore, findings indicated the majority (52.5%) of respondents agreed, and (38.2%) of respondents strongly agreed that the ease to use of the device influenced researchers in their selection. The methods which are commonly used in data storage most of these are easy to use by researchers. The use of desktop computers in the offices, personal laptops, CDs/DVDs, flash memory, and external hard disk in most cases is easy to use by researchers. Furthermore, findings indicated that the majority (47.1%) of respondents agreed that the cost of using the device or methods influenced researchers on the selection of the device/methods to store their data. This implies that the cost of purchasing a data preservation system or device was observed to influence researchers in the selection and use of the method/device. The cost for the initial investment in some research data preservation devices/equipment or technologies is higher. Therefore, researchers' ability to purchase equipment or devices such as computer laptops, CDs/DVDs, flash memory, and external hard disks is cost-effective to researchers this influenced researchers to use such cost effective data preservation devices.

Equally, findings indicated majority (48.5%) of respondents agreed that adequate support provided on the use of data storage/preservation methods/devices influenced researchers to use such devices/methods. The support provided to the use of computer systems, use of flash memory, and other preservation methods influenced researchers in the decision to choose the method to use. For example, SUA has data managers who support researchers in all aspects of research and also the technology transfer office that coordinates research activities. Also, TARI centers have a technology transfer department responsible to coordinate research activities including data preservation and technology transfer. These findings are related to the previous study by Tenopir et al. (2020) who reported that the availability of data managers or data librarians assisted researchers in the procedures how to preserve their research data during, and beyond the project lifecycle.

Furthermore, findings indicated majority (33.8%) of respondents agreed that the availability of technical infrastructure in agriculture research institutions motivates researchers to use such a method/device. This implies that with the presence of institutional or departmental computers, personal laptops it was easy for them to use such available data storage and preservation devices. Findings are in line with prior study by Frederick et al. (2019) revealed that in the selection of data storage infrastructure one should consider the trustworthiness of the infrastructure in the data storage, select a repository that works well with FAIR principles, select a repository that allows granting different access condition to the stored files, and deposit your data to a repository that is accessible to users.

The aspect of reliable power supply in agricultural research institutions shows that (42.6%) of respondents agreed that reliable power supply can influence the use of a certain method/device in data preservation. The use of electronic data preservation devices such as computers, and CDs/DVDs requires a reliable electric supply to a particular research institution. Therefore, the availability of a reliable power supply influenced researchers to store and preserve their data in their desktop office computers, personal laptop computers, CDs, flash sticks, and other electronic devices. The Sokoine University of Agriculture, for example, has standby generators to ensure electricity is available all the time while it is different from some TARI centers where they do not have standby generators this might interfere with data preservation in case of power cut-off.

Recommendations

Based on the study findings the following are the recommendations.

Table 6. Factors that influence the choice of data preservation method/device ($n = 204$).

Factors influence method/device selection	SA		A		U		D		SD	
	F	P	F	P	F	P	F	P	F	P
Device is easy to reach	79	38.7	103	50.3	14	6.9	8	3.9	—	—
Device/method is easy to use	78	38.2	107	52.5	12	5.9	7	3.4	—	—
Device/method is cost effective	51	25	96	47	36	17.6	19	9.3	2	1.0
Adequate support on use of device/method	44	21.6	99	48.5	27	13.2	28	13.7	6	2.9
Adequate data preservation infrastructure	27	13.2	69	33.8	44	21.6	58	28.4	6	2.9
Reliable power supply	33	16.2	87	42.6	25	12.3	52	25.5	7	3.4

Source: Field data, 2021.

SA: strongly agree; A: agree; U: undecided; D: disagree; SD: strongly disagree; F: frequency; P: percentage.

1. It is recommended that agricultural research institutions should establish research data preservation policies and regulations. The policies and regulations should state how agriculture research data will be preserved to ensure availability and accessibility. The policies and regulations for data preservation should be included in the institutional research guidelines.
2. It is recommended that the government should establish an agricultural research data bank to guarantee permanent availability and accessibility of data at all times when they are needed. The modern data preservation methods for example data repositories can be a central area that could enhance data sharing and usage among researchers. Also, the presence of centralized data management could ensure data is preserved permanently for a longer time after the end of the project.

Conclusions

Agriculture research data preservation is imperative in agriculture research institutions for the reason that if research data are preserved well, will promote data access, sharing, and usage. This study intended to investigate data storage, and preservation methods/devices to promote data access and reuse. Findings show that agriculture researchers store and preserve their research data during and beyond the project life cycle. The methods which are prominently used include the use of desktop computers, personal computer (laptops/desktops). This type of data preservation practice is not healthy for the progress of science simply because it does not guarantee data access, sharing, and future reuse. For example, in absence of concerned staff from the station for any reason would deny other researchers to have access and use of some important data. However, storage and preserving data using a personal device is a favorable practice for individual researchers as it permits easy access and use of data during and after the project. To foster the best agriculture research data storage, and preservation practices that could benefit researchers personally and the research institutions, there

is a need to adopt a formal way of preserving data. The presence of an agriculture research data bank can guarantee data availability, and accessibility at all times when they are needed. This study has contributed by proposing the data preservation conceptual framework that has been developed from the variables adopted from the Data Curation Center Model. This conceptual framework will help to enlighten researchers on the need to preserve their data for future use and reuse.

Acknowledgements

Sincere gratitude goes to my study supervisors from the Sokoine University of Agriculture and to my employer Jordan University College for giving me the opportunity to pursue further studies at Sokoine University of Agriculture (SUA).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Nolasko Victory Mwinami  <https://orcid.org/0000-0002-3375-7953>

References

- Adika FO and Kwanya T (2020) Research data management literacy amongst lecturers at Strathmore University Kenya. *Library Management* 41(6/7): 447–466.
- Adrian AM, Emison B, Musker R, et al. (2018) Open access and open data at push universities in partnership with global open data for agriculture (GODAN), (issue June 2018). PUSH University. Available at: <http://wp.auburn.edu/push/wp-content/uploads/2018/06/GODAN-report.pdf> (accessed 23 April 2022).
- Agrawal A (2017) Data roadmaps for sustainable development: Assessment and lessons learned. Available at: <https://www.data4sdgs.org/sites/default/files/servicesfiles/Roadmaps>

- AssessmentReport_Dec2017_final.pdf. (accessed 2 March 2022).
- Aydinoglu A U, Dogan G and Taskin Z (2017) Research data management in Turkey: perceptions and practices. *Library Hi Tech* 35(2): 271–289.
- Barakabitze AA, Kitindi EJ, Sanga C, et al. (2015) New technologies for disseminating and communicating agriculture knowledge and information: Challenges for agricultural research institutes in Tanzania. *The Electronic Journal of Information Systems in Developing Countries* 70(1): 1–22.
- Bhatia V, Stout S, Baldwin B, et al. (2016) Results data initiatives: Findings from Tanzania development gateway. Available at: <https://doi.org/developmentgateway.org/expertise/results> (accessed 15 March 2022).
- Brouder S, Eagle A, Fukagawa NK, et al. (2019) Enabling Open-source Data Networks in Public Agricultural Research. CAST Commentary QTA 2019-1.p.20. <https://www.cast-science.org/wp-content/uploads/2019/05/QTA2019-1-Data-Sharing.p>
- Chidi C (2017) *Investigation of Factors Influencing Data Management in District Health Planning Process: A Case Study of Kibaha Town Council*. Dar-es-salaam: Muhimbili University of Health and Allied Sciences. Available at: <http://dspace.muhas.ac.tz:8080/xmlui/bitstream/handle/123456789/2160/Clara%20Chidi%20%20dissertation%20final.pdf?sequence=1&isAllowed=y> (accessed 21 March 2022).
- Chigwada J, Chiparusha B and Kasiroori J (2017) Research data management in research institutions in Zimbabwe. *Data Science Journal* 16: 31.
- Dileepkumar G (2014) Improving research data management and sharing: Experiences from ICRISAT. In: Proceedings of the winter school on livestock based livelihood options: current status, emerging issues and future scenario in combating agrarian crisis, New Delhi, India, 7 November 2014.
- DPRTC (2019) *Sokoine University of Agriculture Research Regulations and Guidelines*. Sokoine University of Agriculture, 4th edn. Morogoro: Sokoine University of Agriculture. Available at: <https://www.dprtc.sua.ac.tz/wp-content/uploads/SUA-Research-Regulations-and-Guidelines-2019.pdf> (accessed 12 October 2021).
- Eckes AH, Gubala T, Nowakowski P, et al. (2017) Introducing the brassica information portal: Towards integrating genotypic and phenotypic brassica crop data. *F1000 Research* 6: 465–516.
- Elsayed AM and Saleh EI (2018) Research data management and sharing among researchers in Arab universities: An exploratory study. *International Journal of Bioassays* 44(4): 281–299.
- FAO (2017) Country programming framework for united republic of Tanzania 2017–2020, prepared by the government of the united republic of Tanzania, ministry of agriculture, livestock and fisheries in collaboration with the food and agriculture organization of the united nations (FAO). Available at: <https://www.fao.org/3/bt133e/bt133e.pdf> (accessed 26 November 2021).
- Frederick C, Blumzon I and Pănescu AT (2019) *Data Storage*. London: Springer Available at: https://org/10.1007/164_2019_288 (accessed 25 May 2022).
- Funk BYC, Parker K and Calazza T (2018) Women and men in stem at odd over workplace equity: Perceived inequalities are especially common among women in science, technology, engineering and math jobs who work mostly with men. Pew Research Center. Available at: https://www.pewresearch.org/social-trends/wp-content/uploads/sites/3/2018/01/PS_2018.01.09_STEM_final.pdf (accessed 20 May 2022).
- Granda P and Blasczyk E (2011) Data dissemination: Springer reference 688–713. Available at: https://org/10.1007/springerreference_64686 (accessed 5 May 2022).
- Hawkins E, Fulton J, Lee J, et al. (2018) Data fundamentals: Agricultural and biosystems engineering. Technical reports and white papers, p.25. Available at: https://lib.dr.iastate.edu/abe_eng_report/25 (accessed 28 August 2022).
- Higgins S (2008) The DCC curation lifecycle model. *International Journal of Digital Curation* 3(1): 134–140.
- Katabalwa AS, Bates J and Abbott P (2021) Potential opportunities and risks of sharing agricultural research data in Tanzania. *IASSIST Quarterly/International Association for Social Science Information Service and Technology* 45: 3–4.
- Kijazi A (2018) Data availability, exchange and gaps: A case study of Tanzania. Available at: https://library.wmo.int/doc_num.php?explnum_id=5380 (accessed 25 April 2022).
- Kirub A (2016) *Agricultural Research Data Management: Principles, Policy, and Practice*. Addis Ababa: Ethiopian Institute of Agriculture Research. Available at: <http://197.156.72.153:8080/xmlui/bitstream/handle/123456789/217/Data%20management-Book.pdf?sequence=1&isAllowed=y> (accessed 3 February 2022).
- Kruse F and Thestrup J B (2014) Research libraries' new role in research data management, current trends and visions in Denmark. *Liber Quarterly* 23(4): 310–335.
- Maru A (2004) ICT/ICM in agricultural research and development: Status in sub-Saharan Africa. Available at: <https://agris.fao.org/agris-search/search.do?recordID=XR2012210126> (accessed 3 January 2022).
- Mboera LEG (2015) *The Management of Health and Biomedical Data in Tanzania, Need for a National Scientific Data Policy*. Directorate of Information Technology. Dar es Salaam: National Institute for Medical Research. Available at: <https://slideplayer.com/slide/7625850/> (accessed 25 August 2022).
- Methew B and Ross L (2010) *Research Methods: A Practical Guide for the Social Sciences*. London: Longman. Available at: <https://www.pearson.com/uk/educators/higher-education-educators/program/Matthews-Research-Methods-A-Practical-Guide-for-the-Social-Sciences/PGM979381.html> (accessed 27 May 2022).
- Mulder N, Adebamowo CA, Adebamowo SN, et al. (2017) Genomic research data generation, analysis, and sharing – Challenges in the African setting. *Data Science Journal* 16: 1–15.
- Mushi GE, Pienaar H and van Deventer M (2020) Identifying and implementing relevant research data management services for the library at the University of Dodoma, Tanzania. *Data Science Journal* 19: 19.
- Musker R and Schaap B (2018) Global open data in agriculture and nutrition (Godan) initiative partner network analysis [version 1; peer review: 2 approved with reservations]. *F1000 Research* 7(5): 47.
- NBS (2010) *Dissemination and Pricing Policy: The NBS Recognizes That Timely Dissemination of Data to Specialized*

- Users for Research Purposes and Other Decision-Makers is Beneficial*. Dar es Salaam: National Bureau of Statistics. Available at: https://www.nbs.go.tz/nbs/takwimu/ACT/Dissemination_and_Pricing_policy220kb.pdf (accessed 17 March 2022).
- Ng'eno EJ and Mutula S (2018a) Research data management in Kenya's agricultural research institutes. Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (Information Studies) in the School of Social Sciences, College of Humanities, University of KwaZulu-Natal, Pietermaritzburg: South Africa. *Inkanyiso, Journal of Humanities and Social Sciences* 1: 333.
- Ng'eno EJ and Mutula S (2018b) Research data management (RDM) in agricultural research institutes: A literature review, submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (Information Studies) in the School of Social Sciences, College of Humanities, University of KwaZulu-Natal, Pietermaritzburg, South Africa. *Inkanyiso, Journal of Humanities and Social Sciences* 10(1): 28–50.
- Parliament of the United Republic of Tanzania (2016) Tanzania agricultural research institute (TARI) act no 10 of 2016, Dodoma: Tanzania. Available at: https://www.tanzania.go.tz/egov_uploads/documents/sheria_chapa_-Tanzania_Agricultural_Research_inst_sw.pdf (accessed 13 February 2022).
- Plooy-Cilliers A, Vavis C and Bezuidenhout R (2014) *Research Matters*. Cape Town: Juta and Company Ltd.
- Pouchard L (2016) Revisiting the data lifecycle with Big Data curation. *International Journal of Digital Curation* 10(2): 176–192.
- Sielemann K, Hafner A and Pucker B (2020) The reuse of public datasets in the life sciences: Potential risks and rewards. *Peer Journal* 8: e9954.
- Smirnova L, Mergen P, Groom Q J, et al. (2016) Data sharing tools adopted by the European Biodiversity Observation Network Project. *Research Ideas and Outcomes* 2: E9390.
- Tenopir C, Rice NM, Allard S, et al. (2020) Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLoS One* 15(3): e0229003–e0229026.
- TMA (2020) Tanzania meteorological authority (TMA) agrometeorological database. Available at: http://agromet.meteo.go.tz/storage/docs/TMA_Agromet_Database_Manual.pdf (accessed 3 February 2022).
- Tripathi M, Shukla A and Sonkar SK (2017) Research data management practices in university libraries: A study. *DESIDOC Journal of Library & Information Technology* 37(6): 417–424.
- URT (2016) Tanzania Agricultural Research Institute (TARI) act no 10 of 2016. Available at: https://www.tanzania.go.tz/egov_uploads/documents/sheria_chapa_-Tanzania_Agricultural_Research_inst_sw.pdf (accessed 13 February 2022).
- URT (2017) Information communication technology (ICT) policy for national bureau of statistics. Available at: https://www.nbs.go.tz/nbs/takwimu/ACT/ICT_policy_2017.pdf (accessed 22 May 2022).
- Weller T and Monroe-Gulick A (2014) Understanding methodological and disciplinary differences in the data practices of academic researchers. *Library Hi Tech* 32(3): 467–482.
- Yamane T (1967) *Statistics: An Introductory Analysis*, 2nd edn. New York, NY: Harper, and Row.
- Zhao H and Wang J (2015) Development and current situation of agricultural scientific data sharing in China. In: *IFIP Advances in Information and Communication Technology*, vol. 452, pp.80–86. DOI: 10.1007/978-3-319-19620-6_10.

Author biographies

Mr. Nolasko Victory Mwinami is Assistant Lecturer at Jordan University College, Morogoro, Tanzania. Currently is pursuing PhD studies at Sokoine University of Agriculture(SUA), Department of Informatics & Information Technology (DIIT), College of Natural and Applied Sciences (CoNAS), Morogoro, Tanzania.

Prof. Frankwell W. Dulle is Senior Lecturer in the Department of Informatics & Information Technology (DIIT) at Sokoine University of Agriculture in Tanzania. He holds PhD in Information Science, MA (LIS) and BSc (Agric.). He is involved in teaching, research and consultancy activities at Sokoine University of Agriculture in Tanzania

Prof. Wulystan Pius Mtega is Senior Lecture in the Department of Informatics & Information Technology (DIIT) at Sokoine University of Agriculture in Tanzania. He holds PhD in Information Science from the University of South Africa. He is involved in teaching, research and consultancy activities. His areas of interest are knowledge management and Information Communication Technology (ICT) for development. He has written several publications in knowledge management and ICT for development.

CHAPTER THREE PAPER II

JOURNAL OF AGRICULTURAL & FOOD INFORMATION
<https://doi.org/10.1080/10496505.2023.2190899>

 **Routledge**
Taylor & Francis Group

 Check for updates

A Study of Data Sources for Accessibility and Reuse Practices among Agricultural Researchers in Tanzania

Nolasko Victory Mwinami, Frankwell W. Dulle and Wulystan Pius Mtenga

Department of Informatics & Information Technology (DIIT), College of Natural and Applied Sciences (CoNAS), Sokoine University of Agriculture (SUA), Morogoro, Tanzania

ABSTRACT

This paper reports a study that investigated how agricultural researchers use and reuse other researchers' data in Tanzania. This research used a survey method to investigate the factors influencing researchers in this use/reuse. Findings indicated that more than 80% of researchers use and reuse data accessed from different sources. Several factors influence the majority of researchers (more than 70%) to use and reuse data. The results of this study may attract the attention of agricultural researchers elsewhere to agricultural data use and reuse practices.

ARTICLE HISTORY

Received 27 June 2022
 Revised 11 December 2022
 Accepted 10 March 2023

KEYTERMS

Agricultural research; agricultural research data; data reuse; data sharing; data use

Introduction

Accessibility of data is a backbone of scientific research. A greater value is attained for research data when used and reused repeatedly (Shehzad, 2017). When using data acquired from secondary sources such as repositories, reports, and journal articles, researchers are reusing data once collected as primary data (Paff et al., 2019; Sielemann et al., 2020). Research data are used in various ways in scientific research. Researchers use secondary data to compare with new primary data collected, and to review literature for new research projects (Boté & Térmen, 2019; Zhao & Wang, 2015). In this sense, reusing other researchers' data benefits the sciences. Researchers' reuse of data created by others enhances scientific progress (Gregory et al., 2020). Thus, the reuse of scientific data has a potential benefit.

Data sharing is an important activity that enhances data reuse. Data sharing facilitates data accessibility for use and reuse (Late & Kekäläinen, 2020). If agricultural data cannot be located and accessed, these will be of no use even if well gathered (Ali & Dahlhaus, 2022). While scientific

CONTACT Nolasko Victory Mwinami  mwinaminolasko@yahoo.com 

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC

data sharing plays an important role to enhance data reuse, the practice has attracted more attention from researchers, governments, and international organizations and committees. For example, funding institutions, research institutes, the global health community, and international development agencies such as the World Health Organization (WHO), Research Data Alliance, FORCE 11, CODATA, the Wellcome Trust, the Research Council of UK, and the Gates Foundation insist on open access for all funded research data (Adrian et al., 2018; Cooper & Springer, 2019; Khan et al., 2020; Kirub, 2016; Shen, 2015). The accessibility of data ensures data use and reuse by others. The Open Data Initiative is working jointly and struggling to build awareness, teach, and engage data users and publishers about the value of sharing and reusing data (Shehzad, 2017). According to Ali and Dahlhaus (2022), the FAIR principles (Findable, Accessible, Interoperable, and Reusable) emphasize that agricultural data need to be easily accessible for use by others.

The Government of Tanzania took several initiatives to ensure different types of data are accessible for use and reuse. The open data portal, focused on health, education, water, and social-economic data, was established in 2013 for preservation and dissemination of data (Agrawal, 2017). This initiative failed to achieve the targeted objectives; however, some remnants that were left behind gave the public the idea that data need to be open for use by others. In Tanzania, there are different types of statistical data, agricultural administrative data, and climatic and weather data that are shared and accessed via web-based platforms at the National Bureau of Statistics (NBS) and the Tanzania Meteorological Agency (TMA) (Bhatia et al., 2016; Kijazi, 2018; TMA, 2020; United Republic of Tanzania (URT), 2017). Despite the cost of their access, the availability and accessibility of agricultural, weather and climatic data from the NBS and TMA has continued to ensure that data will remain easily accessible for use by different stakeholders including researchers.

These initiatives in Tanzania, aimed at promoting data accessibility through data sharing and reuse, are focused on facilitating public access to governmental data from agencies such as NBS and TMA. Little effort has been made to make agricultural research data, in research institutions, accessible for use among researchers and other stakeholders. Furthermore, due to the absence in Tanzania of a research data sharing policy and a data repository, it is unclear where and how agricultural researchers access their data for reuse (Katabalwa et al., 2021; Mboera, 2015; Tanzania Agriculture Research Institute, 2019). This study seeks to answer the following questions: Where do agricultural researchers find reusable data? How do agricultural researchers reuse other researchers' data? What factors motivate researchers to reuse other researchers' data?

Literature review, theoretical framework, and research model

Agricultural research data

Research data are described as facts, numbers, letters, and symbols that describe an idea; data may be raw, analyzed, observational, experiential, or experimental, upon which an argument is constructed (Omo-Ojugo, 2018; Park, 2018). Agricultural research data can be generated through a range of activities such as animal feeding, husbandry, and crop practices (Wolfert et al., 2017). In the same line, agricultural research data can be generated through the use of agricultural science and technological equipment such as sensors, mobile phones, laboratory tools, satellites, drones, and radar (Wiseman et al., 2019). These scientific data are more valuable when shared and reused for further scientific research.

Data use and reuse

Data reuse is a concept that encompasses various activities, including using your own data for later comparison purposes, using other people's data to compare with newly collected data, and accessing one or more data sets for meta-analysis to address new research questions (Pasquetto et al., 2017). The data reuse process is not a single action that completes the process (Borgman & Pasquetto, 2017). The empirical research on scientific data reuse behaviors can be studied under two categories; the first is through data citation or bibliographic analysis (Park, 2018). In this approach, authors trace data reuse through data citations in literature; for example, searching databases to find journals or books containing the use of authors' data or secondary deployment of original or existing research data to study new questions (Curty, 2015). The second approach is to investigate scientific perceptions, experiences, and attitudes toward data use and reuse in the different research disciplines. This can allow the use of qualitative and quantitative approaches to gather data (Curty, 2015). Data reuse when studied through a qualitative method allows observing where researchers find reusable data, how they evaluate accessed data to fit the intended purpose, how researchers reuse data, what challenges researchers find in reusing data, and what motivates researchers to reuse other people's data (Tenopir et al., 2011). Data reuse also is studied through quantitative methods, such as survey and content analysis, that allow researchers to summarize data reuse experiences and problems and identify the intention to share or reuse data (Curty, 2015; Tenopir et al., 2011). This study applied both qualitative and quantitative approaches to establish agricultural researchers' experiences and attitudes toward data reuse.

Where do researchers find reusable data?

There is a rich literature on data sharing and reuse. Data sharing enables researchers to use other researchers' data. Tenopir et al. (2011) reported that 85% of researchers reused other researchers' data. Such data were accessed from primary sources such as surveys, experiments, and observations, as well as secondary sources, including data accessed from repositories, print publications, museums, CD-ROMS, data consortia, websites, and through personal requests. Wang et al. (2021) used meta-synthesis and inductive coding methods to analyze data reuse in different fields. They found that researchers accessed data beyond their laboratories and their collaborators, discovering data in the secondary sources noted above.

Moreover, a qualitative meta-analysis of two long-term distributed interdisciplinary consortia reveals that researchers accessed data from public collections and fellow researchers (Pasquetto et al., 2019). Researchers used/reused data accessed from local, national, and international networks such as Long-Term Ecological Research Network, Distributed Active-Archive Centers, Global Biodiversity Information Facility, National Biological Information Infrastructure, and National Ecological Observatory Network (Tenopir et al., 2011).

Other studies that focused more on data reuse practices by researchers in the agricultural sciences and related disciplines have shown that agricultural research data and other related information is shared and accessed through various data repositories and from other data infrastructure such as AGRIS, DOAJ (Directory of Open Access Journals), Global Online Research in Agriculture (AGORA), and Brage-UMB (Besemer et al., 2012; Gupta et al., 2016). A recent study conducted in the United States to examine attitudes and opinions of agricultural researchers on open access publishing and data sharing noted that agricultural researchers see the benefits of sharing data via repositories; they also admitted to having used data shared by others (Williams et al., 2019).

In line with this, Williams (2020) explored agricultural data reuse and data sharing based on a review of 102 recent publications by members of the Department of Crop Sciences at the University of Illinois at Urbana-Champaign. Agricultural researchers reused data from a variety of scientific sources including research literature, published articles, and disciplinary repositories (e.g., International Maize and Wheat Improvement Center (CIMMYT), Research Data & Software Repository Network, DDBJ (DNA Data Bank of Japan)).

Most of these previous studies explored data reuse practices by conducting an online survey to review literature or meta-analysis found in online databases to discover data sharing and reuse among researchers; the studies were not specific to agricultural research institutions. The

studies of agricultural research data reuse focused on researchers from one agricultural field, for example, crop sciences (Williams, 2022). Some previous studies used a small sample from which no generalization could be made. The current study employed mixed methods, including a larger sample size to investigate data use and reuse among agricultural researchers from different agricultural fields.

How do agricultural researchers use and reuse other people's data?

Through perusing various literature reviews, it was possible to understand the current data use and reuse practices among researchers. In Wang et al. (2021), the three stages of data reuse are (1) initiation, (2) exploration and collection, and (3) repurposing. Data can be used and reused in different ways. To reuse data effectively data users are required to be expert. Pasquetto et al. (2019) proposed a typology of data reuse ranging from comparative to integrative. Under the comparative purpose data reuse requires expertise in knowing about the quality of data and the ability to assess their value. Data are used for comparative purposes such as ground and calibration, while integrative reuse requires contributory expertise that involves the ability to reuse data in a new experiment. Data integration requires more specialization in scientific knowledge.

Furthermore, researchers use and reuse original or existing data to conduct new analyses, check the quality of newly collected data, study a new research problem, replicate research, validate research, and integrate with other data (Borgman & Pasquetto, 2017; Curty, 2015; Organization for Economic Co-operation and Development (OECD), 2007; Pasquetto et al., 2019; Tenopir et al., 2020). When data are reused for reproducibility purposes, these go through the following process: reanalyzing published data, repeating the study, replicating the findings under different conditions, and reprocessing new data (Pronk, 2019; Wang et al., 2021). A study that examined research data sharing and reuse practices of academic faculty researchers at the Virginia Tech University found that 44% of respondents reported that data can be repurposed and reused by researchers in the same discipline to support new investigations or integrated using new analytical techniques (Shen, 2015). A cross-sectional study to determine the utilization of Health Management Information System (HMIS) data and factors influencing the Tanzanian health system's performance revealed that HMIS data were used to compare performance of health care facilities in terms of service coverage, monitoring of disease trends over time, and providing evidence for community health education and promotion programs (Mboera et al., 2021). Data usage is very important in different research disciplines.

Other studies from agricultural sciences and technology involved examination of attitudes and opinions of agricultural researchers toward open access publishing and data sharing. Williams et al. (2019) indicate that sharing agricultural research data with the public will increase reproducibility and transparency of the research. The adoption of the FAIR data principles in agriculture can improve accessibility of agricultural research data that may lead to improved food security (Ali & Dahlhaus, 2022). Furthermore, agricultural data have been useful in agronomic studies and model inter-comparison, and crop modeling data have been used to accelerate innovation and improve agricultural production (Delserone & Dinkelman, 2016; Paff et al., 2019). However, it remains unclear how agricultural researchers in different fields use and reuse other researchers' data in Tanzania.

What factors influence researchers to reuse other researchers' data?

Joo and Kim (2017) used a survey method to investigate factors that motivated engineering researchers' data reuse behaviors. Findings pointed out that perceived usefulness, perceived concerns, and norms of data reuse motivated researchers to reuse other researchers' data. In addition, Curty (2016) employed a panel discussion to give a clear understanding of factors that facilitate or hinder data reuse in the social sciences. Data reuse behavior was connected to perceived benefits, perceived risks, perceived effort, social influence, facilitating conditions, and intentions to reuse data positively. Through a survey method, Tenopir et al. (2018) investigated motives, attitudes, and data practices of earth and planetary geophysicists; the attitudes of researchers toward data sharing and reuse were the motivators. Using a mixed-methods approach that combined coding with semi-automatic text-searching techniques to assess the impact of data sharing and reuse on data citation in STEM fields, Park (2018) discovered that data sharing and reuse practices were diverse across STEM disciplines.

Other authors have enumerated factors that motivate data sharing and reuse by researchers, including a mandate from funders (Kim & Stanton, 2016; Zenk-Möltgen et al., 2018) and the availability of research funds, which allow such data to be reused (Bezuidenhout & Chakauya, 2018). Data reliability also motivated researchers to reuse other researchers' data (Duan, 2016).

Despite the continuing efforts to promote data reuse, some challenges hinder data reuse practices among researchers. Williams et al. (2019) reveal that participants in an Ithaca S+R study gave the following reasons for not sharing their agricultural research data: the data would not be meaningful to others, lack of money, lack of time, lack of standards for data

sharing in many areas of agricultural research, the data would be difficult to prepare for sharing, and agricultural researchers typically do not receive any reward or recognition in terms of promotion and tenure for sharing their data. In addition to that, lack of data documentation, of accessibility, of specialized knowledge, technical issues, and failure of human scaffolds were identified as challenges to data reuse (Wang et al., 2021).

The literature review found that data reuse practices among researchers in agricultural research institutions in Tanzania are given little attention, as shown by limited documentation of research data from these research institutions. There is a lack of data sharing policy and data repositories in Tanzania (Katabalwa et al., 2021), thus the sources of data used and reused by agricultural researchers from different agriculture fields is unknown. Furthermore, it is not known what drives Tanzanian agricultural researchers to use or reuse other researchers' data.

Theoretical framework and the research model

This study adopted the Theory of Reasoned Action (TRA), which stipulates that a person's intention to perform a given behavior directly determines the performance of actual behavior (Ajzen & Fishbein, 1980). The TRA is a suitable theoretical framework to conceptualize the outcome of attitudes of respondents toward data reuse. The intention is influenced by two factors: the first is an individual's attitude toward behavior (cost and benefits), and the second is the subjective norms referring to the individual's pressure to engage or not to engage in a particular behavior (Ajzen & Fishbein, 1980; Curty et al., 2017; Yzer, 2013).

Subjective norms are important theoretical aspects that shape behavioral intention. This means that the pressure from society may influence a person to engage or not to engage in a particular behavior (Curty et al., 2017). Behavioral intention is a result of both individual influence (attitude toward performing the behavior) and subjective norms. This study adopted facilitating conditions as among the factors that influence a person to perform their actual behavior. Facilitating conditions or availability of infrastructure that preserved data can directly influence researchers to perform the actual behavior of accessing data from a repository and reusing it. Therefore, the presence of accessible data infrastructure may guarantee that a person has direct access to data and the reuse of such data.

The theory helped to measure the key concepts of this study: the dependent variable of the respondent's actual behavior of reusing data or not, and the independent variables, including individuals' perceptions that permit or forbid performing particular behavior (perceived benefits of

data reuse) (Hale et al., 2003), the social perception toward the appropriateness of data reuse (subjective norms) (Nguyen et al., 2018), and the availability of data infrastructure or facilitating conditions that guarantee the availability of data for use/reuse (perform actual behavior). According to Ambarwati et al. (2020), Venkatesh states that facilitating conditions do not affect behavioral intention, but affect actual use behavior. The availability of facilitating conditions such as resources, skills, and technical infrastructure (e.g., data repositories) can play a significant role in the use of the systems (Hamzat & Mabawonku, 2018). In this case, the presence of data platforms may drive an individual's actual data reuse. Almarri et al. (2019) reports that facilitating conditions greatly enhance user satisfaction. Figure 1 shows the conceptual model of factors that influence data use/reuse behavior.

Methods

The current study employed a descriptive and cross-sectional research design to investigate data sharing and reuse among agricultural researchers in Tanzania. The descriptive research survey was performed from February 2020 to March 2022, and enabled the collection and analysis of both qualitative and quantitative data. The qualitative approach was used to observe the circumstances in which researchers find themselves in need of other researchers' data. The quantitative approach was used to identify factors that drive researchers in the intention to reuse other researchers' data. Through the use of the TRA, the factors influencing intention and the actual behavior of data use and reuse were tested to verify the existing theory.

The study was conducted in Tanzania Agricultural Research Institute (TARI) centers and Sokoine University of Agriculture (SUA), as described in Mwinami et al. (2022). From a total population of 527 agricultural

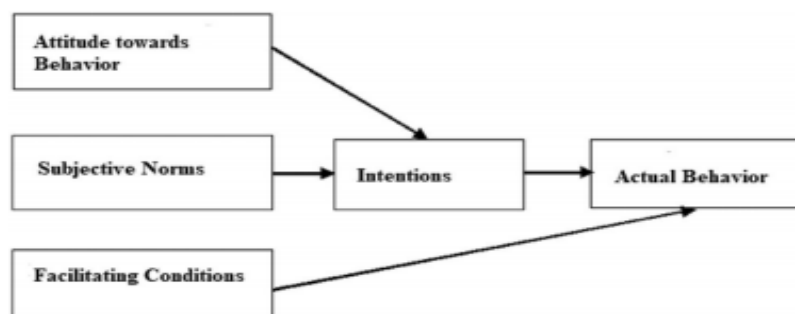


Figure 1. Conceptual model showing factors that influence data use/reuse behavior. Source: Adopted and modified from Ajzen & Fishbein, 1980.

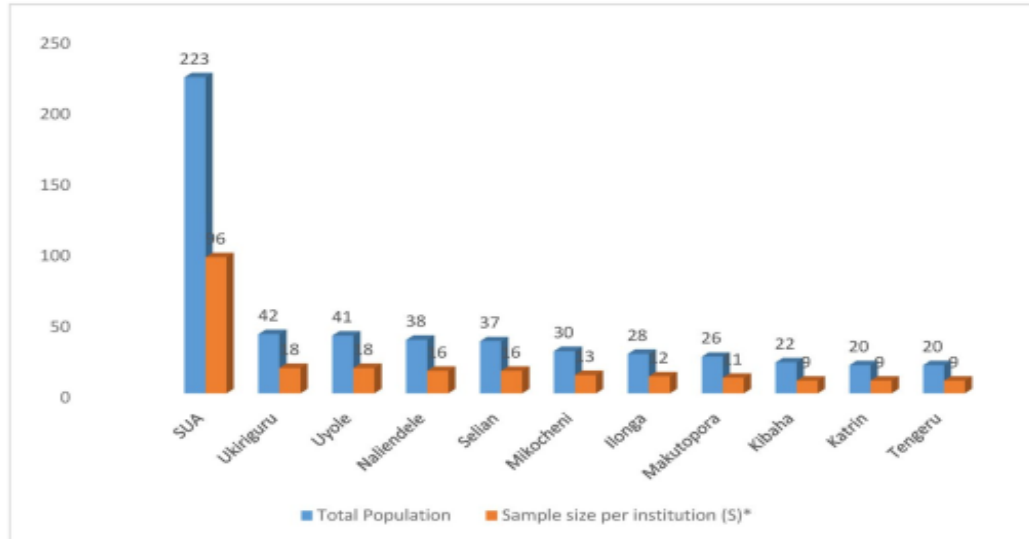


Figure 2. Sampling distribution.

Sample size per institution* = study sample size (227) times population per institution divided by total population.

researchers, Yamane's (1967) formula for calculating sample size was used to obtain the study sample size of 227 (Mwinami et al., 2022). After obtaining the sample size of 227, a representative sample size of respondents per research institution was calculated as shown in Figure 2 (Mwinami et al., 2022).

The survey instrument was tested and found reliable as described in Mwinami et al. (2022). Data collection and analysis were performed as found in Mwinami et al. (2022).

Quantitative data from 204 completed survey questionnaires were analyzed using descriptive statistical techniques that allowed us to obtain frequencies and percentages using Statistical Product and Service and Solution (SPSS) software. On the other hand, qualitative data collected through the focus group discussions and interviews were analyzed using content analysis and arranged in groups with similar subthemes. Qualitative results were used to elaborate on and validate the quantitative findings. Inferential statistics were obtained by running a Binary Logistic Regression technique. This technique was chosen because the dependent variable was a dummy variable (coded 0 = Yes, 1 = No) while the independent variable had several constructs. This technique was used to test the factors that drive actual behavior to use or reuse data. In this model, independent variables, i.e., perceived benefits of data use (attitudes) and subjective norms (society perceptions), and facilitating conditions (availability of data storage infrastructure) were identified to find out how they influence data reuse.

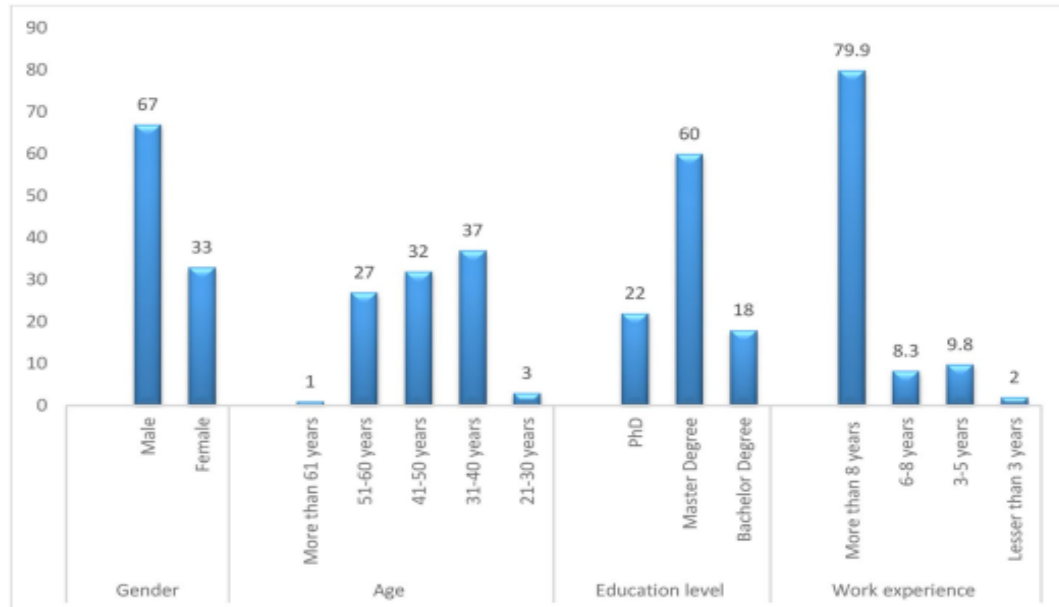


Figure 3. Demographic information of respondents.

Results

Demographic information of respondents

The demographic characteristics of respondents for this survey are presented in [Figure 3](#). The results indicate that a significant percentage (67%) of the researchers who participated in this study were male, implying that the majority of researchers in the surveyed research institutions were male. A significant proportion of researchers (64%) were at an active age (31–60) and engaged in agricultural research. The results indicated that the largest percentage of the respondents (60%) had received a Master’s degree while a few (18%) had completed a Bachelor’s degree. This implies that most of the researchers in the surveyed area had a higher level of education that enabled them to conduct research. A large percentage of the researchers (79.9%) had work experience of more than 8 years. This implies that the majority of the researchers had established research experience and that the junior researchers would be learning and gaining experience from their senior colleagues (Mwinami et al., 2022). Through an interview, one respondent from a TARI center reported:

...some well-experienced researchers have already retired and therefore, it takes time to make those younger researchers well-experienced. It was reported that there is a plan for the Government to bring back retired researchers to work under contracts but I am not sure if it will be implemented. This situation signifies the need to have well-experienced agricultural researchers to achieve the intended research objectives.

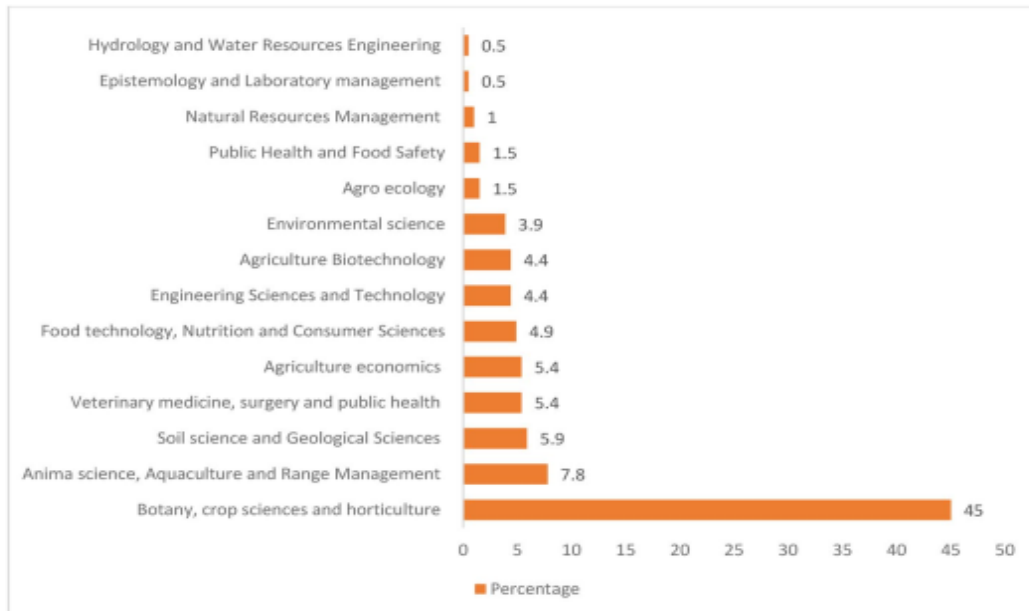


Figure 4. Academic field of specialization of researchers.

Table 1. Sources of agriculture research data used by researchers (N=204).

Data sources	Frequency (n)	Percentage (%)
Data obtained from field surveys	198	97.6
Data obtained from field trials	195	95.6
Requesting data from fellow researchers	193	94.6
Data accessed from other research institutions and universities	189	92.6
Data obtained from research institution library	172	84.3
Data obtained from experiments	165	80.9
Local database (individual, institutional or departmental database)	46	22.5

Researchers' academic field of specialization

The findings in [Figure 4](#) indicate the researchers' academic field of specialization for this survey, with the majority of researchers in botany, crop sciences, and horticulture. The next largest specializations of respondents were in animal science, aquaculture and range management, and agricultural extension and community development.

Where do agricultural researchers access data?

The findings from multiple responses are summarized in [Table 1](#), and indicate that more than 80% of the respondents accessed data from field surveys, field trials, from fellow researchers, from other research centers and universities, from the institutional libraries, or through experimentation. At the focus group discussion at the Ilonga TARI center, researchers reported reusing data from national and international organizations:

It was reported that researchers obtain agricultural research data based on soil, crop improvement, and crop protection through an observation process. Also, other different types of agricultural research data based on the climate and weather are accessed from the Tanzania Meteorological Agency (TMA) and the National Bureau of Statistics (NBS) website and database. We obtain different types of data from other research collaborators and funders, including the International Institute of Tropical Agriculture (IITA), International Research Institute for Climate and Society (IRI), International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Centro Internacional de Agricultura Tropical (CIAT), Technical Centre for Agriculture and Rural Cooperation (CTA), the Food and Agriculture Organization (FAO), Access to Global Online Research (AGORA) and the International Maize and Wheat Improvement Centre (CIMMYT).

An interview with one of the key informants revealed the following:

...research data are generated through laboratory experiments, observation, and field trials. I access data from secondary sources such as local and international online data sets (NBS, FAOSTAT) and other databases or platforms owned by project collaborators/partners. Also, I reuse secondary data by accessing them from the institution library when reviewing the literature. I get data through requesting from my fellow researchers, and communicating with other research institutions that are found within and outside our country.

Inferential statistics for sources of data for reuse across research institutions

Results from cross-tabulations of data sources and reuse (Table 2) show that there are variations in the researchers' sources of agricultural data across the institutions, with field surveys, trials, experiments, fellow researchers, and other research institutions being used predominantly over local or individual databases. The Ilonga, Naliendele, and Selian institutions did not report using local databases as a source of agricultural research data. Moreover, the use of the library, local databases, and experiments as a source of agricultural research data were significant ($p \leq 0.005$).

Type of research data accessed and reused

As shown in Table 3, findings indicate that 97.1% of respondents accessed and used both primary and secondary data, while 2.5% said that they accessed and used secondary data. Only 0.5% of respondents said they only accessed and used primary data.

Rate of researchers' use/reuse of other researchers' data

Results found in Table 4 indicate that about 51% always reused research data when writing research project proposals, papers, and technical reports,

Table 2. Inferential statistics for sources of data for reuse across research institutions (N=204).

Source of data	SUA		Ilonga		Naliendele		Selian		Uyole		Tengenu		Katrin		Ukiriguru		Makutopora		Mikocheni		Kibaha		p-value	
	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%		
Institutional library	76	42.2	11	6.4	15	8.7	5	2.9	16	9.3	8	4.7	8	4.7	9	5.2	8	4.7	7	4.1	9	5.2	0.001	
Fellow researchers	74	39.2	9	4.8	16	8.5	14	7.4	16	8.5	8	4.2	8	4.2	14	7.2	10	5.3	11	5.8	9	4.8	0.162	
Field survey	80	40.6	11	5.6	16	8.1	15	7.6	16	8.1	8	4.1	6	3	15	7.6	10	5.1	11	5.6	9	4.9	0.097	
Local databases	38	82.6	0	0	0	0	0	0	2	4.3	2	4.3	3	6.5	0	0	0	0	0	0	0	1	2.1	0.001
Experiments	70	42.9	8	4.9	13	8	7	4.3	14	8.6	7	4.3	8	4.9	15	9.2	9	5.5	6	3.7	6	3.7	0.005	
Field trials	81	41.5	10	5.1	16	8.2	12	6.2	12	6.2	8	4.1	8	4.1	15	7.7	10	5.1	11	5.6	8	4.1	0.057	
Other Tanzania Agricultural Research Institutions	71	38.1	12	6.4	16	8.6	14	7.5	8	4.3	7	3.7	7	3.7	15	8	10	5.3	10	5.3	8	4.3	0.227	

F= Frequency (n)

Table 3. Type of research data reused by researchers (N=204).

Type of research data accessed	Frequency	Percentage (%)
Both primary and secondary agricultural research data	198	97.1
Secondary agricultural research data	5	2.5
Primary agricultural research data	1	0.5
Total	204	100

Table 4. Rate of researchers reuse of others researchers' data (N=204).

Rate of usage	Frequency (n)	Percentage (%)
Always	104	51.0
Occasionally	61	29.9
Very rarely	38	18.6
Never	1	0.5
Total	204	100

Table 5. How data are used/reused by researchers (N=204).

Reuse of agriculture research data	Frequency (n)	Percentage (%)
Used in writing scientific paper	194	95.1
Answer scientific question	177	86.8
Aid interpretation of results	175	85.8
Used in reproducing new study	173	84.8
Used in comparing results	169	82.8
Used in shaping research question	165	80.9
Used as reference in training and teaching	163	79.9
Used in testing theories	157	77.0
Used in paper review (meta-analysis)	152	74.5

while about 30% occasionally did so. However, about 18.6% of respondents said that they reused research data very rarely.

How do researchers reuse other researchers' data?

More than 80% of the respondents indicated that they use/reuse agricultural research data in writing scientific papers, answering scientific questions, to aid interpretation of results, reproducing new studies, comparing results, and using data in shaping research questions (Table 5). A key informant stated:

Researchers from research and academic institutions at SUA use and reuse data as reference materials during teaching and training students in various agriculture fields. Also, research data are used when writing research proposals for requesting project funds. Also, I use research data when writing review papers or conducting meta-analyses.

Another interviewee reported:

I use data while preparing research technical reports (annual reports) to be submitted to funders/collaborators, and TARI Headquarters. Research data are facts, or any innovation is used when providing technical advice and training rural communities about new technologies generated through research.

Table 6. Factors influencing researchers to use/reuse other researchers' data (N=204).

Factors influence	Frequency (n)	Percentage (%)
Attitude		
Data reuse improves my results due to similarity in study purposes	164	80.4
Data reuse helps me answer my study questions	154	75.5
Data reuse saves time for researcher	152	74.5
Data reuse is efficient due the use of similar tools in new study	135	66.2
Data reuse saves money/finance	133	65.2
Subjective Norms		
Institutional research arrangement	114	56.4
Funders' or project collaborators' policies	108	53.4
Facilitating conditions		
Quality data from trusted data storage infrastructure	154	74.5
Reliable data from trusted data storage infrastructure	163	79.9
Accessible data from trusted data storage infrastructure	129	63.2

Factors influencing researchers to reuse other researchers' data

The results presented in Table 6 indicate that more than 70% of the respondents were influenced by several factors to reuse other researchers' data. Based on the TRA, the factors included attitudes, subjective norms, and facilitating conditions. These factors are: data reuse improves results due to similarity in study purpose (attitude), data reused were those from trusted storage infrastructure (facilitating conditions), data were reused because they help answer research questions (attitude), data reuse saves time (attitude), and data were reused because they had good quality (attitude).

Modeling factors inducing researchers to reuse other researchers' data

The Binary Logistic Regression Model was used in testing the relationship between dependent and independent variables. The results found in Table 7 indicate that there is a relationship between the dependent variable (data use/reuse) and independent variables (attitude, subjective norms, and facilitating conditions). However, none of the independent variables had a statistically significant difference, inferring that all of them had an equal influence on data use and reuse (actual behavior).

Discussion

Demographic information of respondents

Results imply that male agricultural researchers superseded female researchers within the institutions under study. This situation has been a tradition in many agricultural research institutes and in universities that have specialization in agriculture. It is not surprising to find more male employees than female employees in this field. It is also worth noting that all of these institutions are dealing with scientific inventions

Table 7. Modeling factors influencing researchers to reuse other researchers' data.

Factors	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
							Lower	Upper
Facilitating conditions								
Reliable data from trusted storage	-.910	.888	1.050	1	.305	.403	.071	2.293
Accessible data from trusted storage	-.342	.618	.305	1	.581	.711	.211	2.387
Quality data from trusted storage	.283	.635	.198	1	.656	1.327	.382	4.608
Attitude								
Data reuse saves money	-.111	.658	.029	1	.866	.895	.246	3.250
Data reuse improved my results	-.042	.825	.003	1	.959	.958	.190	4.827
Data reuse is efficient	-.320	.735	.189	1	.663	.726	.172	3.068
Data reuse helps to answer study question	.203	.757	.072	1	.788	1.225	.278	5.401
Data reuse saves time	.568	.729	.607	1	.436	1.765	.423	7.369
Subjective Norms								
Institutional research arrangement	-.148	.603	.060	1	.807	.863	.265	2.813
Funders' policies	-.080	.619	.017	1	.897	.923	.274	3.102
Constant	3.082	.968	10.133	1	.001	21.806		

and discoveries whose implementations require a background in science subjects. Traditionally in Tanzania and neighboring countries, physical science subjects have been learnt by more males than females. The findings of the current study are related to those of the study by Funk et al. (2018), which reveals that fewer women were working in jobs related to the physical sciences, technology, engineering and mathematics (STEM).

In terms of age group distribution, this study found that the majority of the researchers in agricultural research institutions were aged between 31 and 50 years and few were between the age groups of 21–30 years. Only two respondents who participated in this study were more than 61 years old. From this we can assume that researchers who were aged 31 years and above are experienced agricultural researchers participating in different research projects at SUA and the TARI centers. These results imply that researchers aged 31 years or above have been working with research projects for many years and therefore have generated agricultural research data that may be used by other researchers. Furthermore, researchers at the mentioned age have been using and reusing research data they accessed from other researchers or other sources.

Regarding the respondents' educational levels, the findings indicate that agricultural researchers are adequately qualified for their work as 82% had a Master's degree and above in the agricultural sciences. The respondents with this academic qualification have enough knowledge and skills to conduct research activities in the different agricultural fields as per their institutions' mandate and research priorities. In terms of experience, the results signify that agricultural research activities require researchers who are adequately well-experienced to achieve the intended goal of ensuring food security is achieved through improving seeds and farming methods.

Where do agricultural researchers find data to use and reuse?

Our findings indicate that agricultural researchers access and reuse both primary and secondary agricultural research data. They use their own primary agricultural research data or raw data obtained through research activities such as field surveys, field trials, observations, and laboratory experiments. The secondary data are from secondary sources such as repositories, literature reviews, and reading previous research reports. These results are in line with the study by Curty (2015), in which primary research data are used when a researcher collects data for a specific project and then uses the same data later on for the same or a later project. When datasets are contributed to a repository and retrieved by someone else and reused for a new project, it would be considered secondary data analysis.

Furthermore, a study conducted by Park (2018) details several types of data that were reused by STEM researchers in astronomy/physics and chemistry (crystal structure, crystallographic data, and molecular data); in computing (software, code, and models); in the earth sciences (datasets); in engineering (test data, datasets, and GIS vector data); in mathematical sciences (software, Matrix); and in technology (datasets, file sets, and TIFF images). Furthermore, data that are reused are those obtained from sources such as remote sensing, weather forecast services, and genome portals that provide resources for biotechnologists (Alliance for Internet of Things and Edge Computing Innovation (AIOTI), 2020; Boté & Térmens, 2019; Seol et al., 2016). According to key informant interviews, researchers use/reuse other researchers' data via local and international data sources such as FAOSTAT and ICRISAT. Researchers also reuse data that are accessed from other institutions, collaborators, partners, and funder's datasets.

The inferential statistical results indicate that the use of the library, local databases, and experiments had statistical significance, implying that these institutions vary in the way they obtain data from these sources. The institutions differ in the extent to which they conduct experimental research, other than SUA, a research institution. The latter, on the other hand, has a library with an institutional repository and subscribes to many e-databases that offer researchers access to agricultural research data with opportunities to conduct secondary data analysis or meta-analysis.

How researchers use/reuse agricultural research data

Researchers use and reuse data in writing scientific papers; for example, researchers use their original research data or raw data to answer intended research questions. Researchers reuse other researchers' data by reviewing different literature or doing meta-analysis to compare results or make critical analyses. Furthermore, researchers reuse scientific data in writing

scientific papers, technical reports, and in project proposal preparation. These findings are related to those that reveal that research data were reused in writing research papers and dissertations (Curty, 2015; Kadir & Yunus, 2017). Our findings show that researchers reuse available data to compare results with what has been obtained in their current study. These findings are related to those that examine data reuse as an act of returning to one's data for later comparison purposes (Pasquetto et al., 2019), and to another in which agricultural researchers in research institutes generate, collect, store and use data for research and decision-making in agriculture (Ng'eno & Mutula, 2018).

Agricultural researchers use data as a reference in teaching and learning, especially in research institutions where there are some academic activities. These findings are related to an earlier study, which revealed data accessed from the UKDA data bank facilitated secondary analysis and teaching and learning (Kadir & Yunus, 2017). It has been shown that data are used in shaping and answering intended research problems and facilitating the interpretation of results, aligning with previous studies which reveal that research data are collected for different reasons, including pursuing new research projects and answering a variety of research questions within the study and generating new study hypotheses (Linek et al., 2017; Pouchard, 2015; Zuiderwijk et al., 2020).

Factors influencing researchers to reuse other researchers' data

In terms of factors causing researchers to reuse others' data, respondents reported several factors that motivate them. Such factors include the quality of data available from different sources, perceived benefits, and availability of data storage infrastructure. Findings obtained from Binary Logic Regression Model (Table 7) indicate no independent variables had a statistical significance, meaning that all the factors contributed equally to the actual behavior of data use and reuse. The reuse of existing data that are similar to a study's purpose were observed to drive researchers' intention to use that particular data. These findings are in line with the study by Joo and Kim (2017), in which researchers reused existing data that were seen to merit a new study.

Good quality data were observed to influence researchers in data reuse, as was the case with social scientists who reused data of good quality in terms of completeness and clarity (Curty et al., 2017; Zuiderwijk et al., 2020). The data accessible from trusted reliable sources or infrastructure was observed to induce researchers to reuse such data, as found in earlier studies (Caracciolo & Keizer, 2011; Koopman & De Jager, 2016). Saving time and minimizing cost were among the factors that drive researchers to reuse other researchers' data, matching previous reports that researchers

prefer to use data that are obtained or accessed by fellow researchers to save time and minimize the cost and duplication of efforts (Pronk, 2019; Sielemann et al., 2020).

Our findings also reveal that agricultural researchers reuse data that are openly published. Funding agencies insist that all data resulting from publicly funded research projects should be made open for reuse by others. This policy is beneficial to researchers because they can reuse other researchers' resources when writing a proposal, report or conducting meta-analysis. These findings align with studies that reveal that researchers were required to share their data so that they may be reused by others (Bezuidenhout & Chakauya, 2018; Delserone & Dinkelman, 2016; Zenk-Möltgen et al., 2018). During the interviews, researchers report that they access data from fellow researchers. Internal arrangements in research institutions also motivate researchers to reuse colleagues' data. Researchers working in teams with other project collaborators and funders allow them to exchange resources, which enables data sharing and reuse.

Conclusion

It can be concluded that agricultural researchers reuse data that are accessed from primary and secondary sources. There are factors that influence agricultural researchers to use and reuse other researchers' data. The use and reuse of existing accessible agricultural research data has a positive impact on the improvement of agricultural research that may facilitate the attainment of development goals by ensuring food security in the country.

Recommendations

1. Agricultural researchers should put more effort into exploiting their existing data and using original data during their research activities so that they achieve the best in all aspects of research in different agricultural fields.
2. In order to facilitate future use/reuse of data generated by researchers from different research projects, it is recommended that research institutions should establish their own local databases for hosting all agricultural research data.
3. Agricultural research institutions should formulate and implement a data sharing policy and establish standards or tools that may facilitate data reuse.
4. Researchers should be willing to allow their data to be reused by others simply because they too admit using or reusing the data that are accessed from other sources.

Acknowledgments

Sincere gratitude goes to my study supervisors from the Sokoine University of Agriculture, and to my employer - Jordan University College - for permitting me to pursue further studies. My sincere gratitude also goes to the Tanzania Commission for Science and Technology (COSTECH) for allowing me to attend the scientific paper writing training that added value to my skills and knowledge in writing this scientific paper.

References

- Adrian, A. M., Emison, B., Musker, R., & Hopkinson, J. (2018). *Open access & open data at PUSH universities in partnership with Global Open Data for Agriculture (GODAN)*. Auburn University. <http://wp.auburn.edu/push/wp-content/uploads/2018/06/GODAN-report.pdf>
- Agrawal, A. (2017, December 18). *Data roadmaps for sustainable development: Assessment and lessons learned*. Global Partnership for Sustainable Development Data. <https://www.data4sdgs.org/resources/data-roadmaps-sustainable-development-assessment-and-lessons-learned>
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behaviour*. Prentice-Hall.
- Ali, B., & Dahlhaus, P. (2022). The role of FAIR data towards sustainable agricultural performance: A systematic literature review. *Agriculture*, 12(2), 309. <https://doi.org/10.3390/agriculture12020309>
- Alliance for Internet of Things and Edge Computing Innovation (AIOTI). (2020, 10 June). *Data sharing in agriculture: Towards a European agriculture data space [Webinar]*. Alliance for Internet of Things and Edge Computing Innovation (AIOTI). https://aioti.eu/wp-content/uploads/2020/07/Report_Data_Sharing_in_Agriculture_Online_Webinar_10.06.2020_Final.pdf
- Almarri, H., Ameen, A., Isaac, O., Khalifa, G. S. A., & Bhaumik, A. (2019). Antecedents and outcomes of online social networks (OSN) usage among public sector employees. *International Journal of Psychosocial Rehabilitation*, 24(6), 6373–6388. <https://doi.org/10.37200/V2416/18233>
- Ambarwati, R., Harja, Y. D., & Thamrin, S. (2020). The role of facilitating conditions and user habits: A case of Indonesian online learning platform. *The Journal of Asian Finance, Economics and Business*, 7(10), 481–489. <https://doi.org/10.13106/jafeb.2020.vol7.no10.481>
- Besemer, H., Edge, P., Subirats, I., Tsiflidou, E., & Protonotarios, V. (2012). *Review of existing agricultural data management practices, lifecycles, and workflows*. AgINFRA. <https://www.fao.org/3/aq190e/aq190e.pdf>
- Bezuidenhout, L., & Chakauya, E. (2018). Hidden concerns of sharing research data by low/middle-income country scientists. *Global Bioethics = Problemi di Bioetica*, 29(1), 39–54. <https://doi.org/10.1080/11287462.2018.1441780>
- Bhatia, V., Stout, S., Baldwin, B., & Homer, D. (2016). *Results data initiatives: Findings from Tanzania*. Development Gateway. <https://developmentgateway.org/wp-content/uploads/2020/10/RDI-Tanzania.pdf>
- Borgman, C., & Pasquetto, I. (2017). *The Big Data to Knowledge (BD2K) guide to the fundamentals of data science: Why data sharing and reuse are hard to do*. UCLA Center for Knowledge Infrastructures. https://escholarship.org/content/qt0jj17309/qt0jj17309_noSplash_a9f596094db9852e98e516345b3efc33.pdf?t=p7agej
- Boté, J.-J., & Térmens, M. (2019). Reusing data: Technical and ethical challenges. *DESIDOC Journal of Library & Information Technology*, 39(6), 329–337. <https://doi.org/10.14429/djlit.39.06.14807>

- Caracciolo, C., & Keizer, J. (2011). *A framework for knowledge sharing and interoperability in agricultural research for development*. Food and Agriculture Organization of the United Nations (FAO). <https://www.fao.org/3/an651e/an651e00.pdf>
- Cooper, D., & Springer, R. (2019, 13 May). Data communities: A new model for supporting STEM data sharing [issue brief]. Ithaca S + R. <https://digitalcommons.unl.edu/scholcom/109/>
- Curty, R. G. (2015). *Beyond "data thrifting": An investigation of factors influencing research data reuse in the social sciences*. [Doctoral Dissertation. Syracuse University]. Surface. <https://surface.syr.edu/etd/266>
- Curty, R. G. (2016). Factors influencing research data reuse in the social sciences: An exploratory study. *International Journal of Digital Curation*, 11(1), 96–117. <https://doi.org/10.2218/ijdc.v11i1.401>
- Curty, R. G., Crowston, K., Specht, A., Grant, B. W., & Dalton, E. D. (2017). Attitudes and norms affecting scientists' data reuse. *PLoS One*, 12(12), e0189288. <https://doi.org/10.1371/journal.pone.0189288>
- Delserone, L. M., & Dinkelman, A. L. (2016). *Investigating the practices and needs of agricultural researchers at the University of Nebraska-Lincoln*. Ithaca S + R and the UNL Libraries. <https://digitalcommons.unl.edu/librarianscience/341/>
- Duan, P. (2016). Essay on the external communication strategies of the China Dream: Analysis and study of reports on the China Dream in the Washington Post and on CNN. *Global Media and China*, 1(3), 261–270. <https://doi.org/10.1177/205943641667146>
- Funk, C., Parker, K., & Caiazza, T. (2018). *Women and men in STEM often at odds over workplace equity*. Pew Research Centre. <http://hdl.handle.net/10919/92671>
- Gregory, K., Groth, P., Scharnhorst, A., & Wyatt, S. (2020). Lost or found? Discovering data needed for research. *Harvard Data Science Review*, 2(2). <https://doi.org/10.1162/99608f92.e38165eb>
- Gupta, N. K., Isaac, R. K., & Singh, R. K. (2016). Maintenance and analysis of agricultural data: A challenge. *International Journal of Bioassays*, 5(9), 4842–4848. <https://doi.org/10.21746/ijbio.2016.09.004>
- Hale, J. L., Householder, B. J., & Green, K. L. (2003). The theory of reasoned action. In *The persuasion handbook: Developments in theory and practice*. Sage. <https://doi.org/10.4135/9781412976046.n14>
- Hamzat, S., & Mabawonku, I. (2018). Influence of performance expectancy and facilitating conditions on use of the digital library by engineering lecturers in universities in South-west, Nigeria. *Library Philosophy and Practice (e-Journal)*, 1670. <https://digitalcommons.unl.edu/libphilprac/1670/>
- Joo, Y., & Kim, Y. (2017). Engineering researchers' data reuse behaviours: A structural equation modelling approach. *The Electronic Library*, 35(6), 1141–1161. <https://doi.org/10.1108/EL-08-2016-0163>
- Kadir, I. K. A., & Yunus, A. M. (2017). Recent projects of a preservation framework for digital preservation. *International Journal of Academic Research in Business and Social Science*, 7(11), 16–38. <https://doi.org/10.6007/ijarbss/v7-i11/3431>
- Katabalwa, A. S., Bates, J., & Abbott, P. (2021). Potential opportunities and risks of sharing agricultural research data in Tanzania. *IASSIST Quarterly*, 45(3-4). <https://doi.org/10.29173/iq997>
- Khan, N., Pink, C. J., & Thelwall, M. (2020). Identifying data sharing and reuse with Scholix: Potentials and limitations. *Patterns*, 1(1), 100007. <https://doi.org/10.1016/j.patter.2020.100007>
- Kijazi, A. (2018, 11–12 October). *Data availability, exchange and gaps: A case study of Tanzania*. Tanzania Meteorological Agency. https://library.wmo.int/doc_num.php?explnum_id=5380

- Kim, Y., & Stanton, J. M. (2016). Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, 67(4), 776–799. <https://doi.org/10.1002/asi.23424>
- Kirub, A. (2016). *Agricultural research data management: Principles, policy, and practice*. Ethiopian Institute of Agriculture Research. https://www.researchgate.net/publication/308030253_Agricultural_Research_Data_Management_PRINCIPLES_POLICY_and_PRACTICE
- Koopman, M. M., & De Jager, K. (2016). Archiving South African digital research data: How ready are we? *South African Journal of Science*, 112(7/8), 7. <https://doi.org/10.17159/sajs.2016/20150316>
- Late, E., & Kekäläinen, J. (2020). Use and users of a social science research data archive. *PLoS One*, 15(8), e0233455. <https://doi.org/10.1371/journal.pone.0233455>
- Linek, S. B., Fecher, B., Friesike, S., & Hebing, M. (2017). Data sharing as a social dilemma: Influence of the researcher's personality. *PLoS One*, 12(8), e0183216. <https://doi.org/10.1371/journal.pone.0183216>
- Mboera, L. E. G. (2015). *The management of health and biomedical data in Tanzania: Need for a national scientific data policy*. National Institute for Medical Research. <https://slideplayer.com/slide/7625850/>
- Mboera, L. E. G., Rumisha, S. F., Mbata, D., Mremi, I. R., Lyimo, E. P., & Joachim, C. (2021). Data utilization and factors influencing the performances of the health management information systems in Tanzania. *BMC Health Services Research*, 21(1), 498. <https://doi.org/10.1186/s12913-021-06559-1>
- Mwinami, N. V., Dulle, F. W., & Mtega, W. P. (2022). Data preservation practices for enhancing agricultural research data usage among agricultural researchers in Tanzania. *Journal of Librarianship and Information Science*, 0(0) <https://doi.org/10.1177/09610006221138110>
- Ng'eno, E., & Mutula, S. (2018). Research data management (RDM) in agricultural research institutes : A literature review. *Inkanyiso: Journal of Humanities and Social Sciences*, 10(1). <https://www.ajol.info/index.php/ijhss/article/view/183277>
- Nguyen, Q. A., Hens, L., MacAlister, C., Johnson, L., Lebel, B., Tan, S. B., Nguyen, H. M., Nguyen, T. N., & Lebel, L. (2018). Theory of reasoned action as a framework for communicating climate risk: A case study of schoolchildren in the Mekong Delta in Vietnam. *Sustainability*, 10(6), 2019. <https://doi.org/10.3390/su10062019>
- Omo-Ojugo, E. (2018). Relevance of big data analytics in agriculture: Focus on Nigeria agricultural sector. *International Journal of Science and Research (IJSR)*, 7(9), 1–11. https://www.ijsr.net/get_abstract.php?paper_id=ART2019910
- Organization for Economic Co-operation and Development (OECD). (2007, 12 April). *OECD principles and guidelines for access to research data from public funding*. OECD. https://www.oecd-ilibrary.org/science-and-technology/oecd-principles-and-guidelines-for-access-to-research-data-from-public-funding_9789264034020-en-fr
- Paff, K. E., Assey, S., Araya, A., Davison, J., Getu, A. A., & Giday, O. (2019). Experimental field data for modelling the growth response of tef to nitrogen fertilizer and water stress. *Open Data Journal for Agricultural Research*, 6, 8–13. <https://doi.org/10.18174/odjar.v5i0.16228>
- Park, H. (2018). *The impact of research data sharing and reuse on data citation in STEM Fields*. [Doctoral Dissertation. University of Wisconsin Milwaukee]. UWM Digital Commons. <https://dc.uwm.edu/etd/2005>.
- Pasquetto, I. V., Borgman, C. L., & Wofford, M. F. (2019). Uses and reuses of scientific data: The data creators advantage. *Harvard Data Science Review*, 1(2). <https://doi.org/10.1162/99608f92.fc14bf2d>

- Pasquetto, I. V., Randles, B. M., & Borgman, C. L. (2017). On the reuse of scientific data. *Data Science Journal*, 16, 8. <https://doi.org/10.5334/dsj-2017-008>
- Pouchard, L. (2015). Revisiting the data lifecycle with big data curation. *International Journal of Digital Curation*, 10(2), 176–192. <https://doi.org/10.2218/ijdc.v10i2.342>
- Pronk, T. E. (2019). The time efficiency gain in sharing and reuse of research data. *Data Science Journal*, 18(1), 10. <https://doi.org/10.5334/dsj-2019-010>
- Seol, Y.-J., Lee, T.-H., Park, D.-S., & Kim, C.-K. (2016). NABIC : A new access portal to search, visualize and share agricultural genomics data. *Evolutionary Bioinformatics*, 12, EBO.S34493. <https://doi.org/10.4137/EBO.S34493>
- Sielemann, K., Hafner, A., & Pucker, B. (2020). The reuse of public datasets in the life sciences: Potential risks and rewards. *PeerJ*, 8, e9954. <https://doi.org/10.7717/peerj.9954>
- Shehzad, M. (2017). *Open Data Initiative: Understanding management in an uncertain ecosystem* [Unpublished master's thesis]. UMEA Universitet. <http://umu.diva-portal.org/smash/get/diva2:1149217/fulltext01.pdf>
- Shen, Y. (2015). Research data sharing and reuse practices of academic faculty researchers: A study of the Virginia Tech data landscape. *International Journal of Digital Curation*, 10(2), 157–175. <https://doi.org/10.2218/ijdc.v10i2.359>
- Tanzania Agricultural Research Institute. (TARI). (2019). *Communication strategy*. Tanzania Agricultural Research Institute. <https://www.tari.go.tz/publications/details/communication-strategy->
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS One*, 6(6), e21101. <https://doi.org/10.1371/journal.pone.0021101>
- Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Research data sharing: practices and attitudes of geophysicists. *Earth and Space Science*, 5(12), 891–902. <https://doi.org/10.1029/2018EA000461>
- Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLoS One*, 15(3), e0229003. <https://doi.org/10.1371/journal.pone.0229003>
- TMA. (2020). *Agrometeorological Database*. Tanzania Meteorological Authority. http://agromet.meteo.go.tz/storage/docs/TMA_Agromet_Database_Manual.pdf
- United Republic of Tanzania (URT). (2017). *ICT policy for National Bureau of Statistics*. National Bureau of Statistics. https://www.nbs.go.tz/nbs/takwimu/ACT/ICT_POLICY_2017.pdf
- Wang, X., Duan, Q., & Liang, M. (2021). Understanding the process of data reuse: An extensive review. *Journal of the Association for Information Science and Technology*, 72(9), 1161–1182. <https://doi.org/10.1002/asi-24483>
- Williams, S. C., Farrell, S. L., Kerby, E. E., & Kocher, M. (2019). Agricultural researchers' attitudes toward open access and data sharing. *Issues in Science and Technology Librarianship (ISTL)*, 91. <https://doi.org/10.29173/istl4>
- Williams, S. C. (2022). Data practices ten years later: A new review of selected publications by crop sciences faculty. *Journal of Agricultural & Food Information*, 23(1-2), 28–43. <https://doi.org/10.1080/10496505.2021.2013850>
- Wiseman, L., Sanderson, J., Zhang, A., & Jakku, E. (2019). Farmers and their data: An examination of farmers' reluctance to share their data through the lens of the laws impacting smart farming. *NJAS: Wageningen Journal of Life Sciences*, 90-91(1), 1–10. <https://doi.org/10.1016/j.njas.2019.04.007>
- Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming, A review. *Agricultural Systems*, 153, 69–80. <https://doi.org/10.1016/j.agry.2017.01.023>
- Yamane, T. (1967). *Statistics: An introductory analysis*. (2nd ed.). Harper and Row.

- Yzer, M. (2013). Reasoned action theory: Persuasion as belief-based behavior change. In *The SAGE handbook of persuasion: Developments in theory and practice* (pp.120–136). Sage Publications. <https://doi.org/10.4135/9781452218410>
- Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V., & Balaban, E. (2018). Factors influencing the data sharing behavior of researchers in sociology and political sciences. *Journal of Documentation*, 74(5), 1053–1073. <https://doi.org/10.1108/JD-09-2017-0126>
- Zhao, H., & Wang, J. (2015). Development and current situation of agricultural scientific data sharing in China. In D. Li & Y. Chen (Eds.), *Computer and computing technologies in agriculture VIII, IFIP advances in information and communication technology* (Vol. 452, pp. 80–86). Springer. https://doi.org/10.1007/978-3-319-19620-6_10
- Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLoS One*, 15(9), e0239283. <https://doi.org/10.1371/journal.pone.0239283>

CHAPTER FOUR

PAPER III



Vol. 33, (2023) 1–28 | e-ISSN: 2213-056X

Communication Channels and their Potential Applicability in Enhancing Agricultural Research Data Sharing among Agricultural Researchers in Tanzania

Nolasko Victory Mwinami

Department of Informatics & Information Technology, College of Natural and Applied Sciences, Sokoine University of Agriculture, Morogoro, Tanzania, mwinaminolasko@yahoo.com, orcid.org/0000-0002-3375-7953

Frankwell W. Dulle

Department of Informatics & Information Technology, College of Natural and Applied Sciences, Sokoine University of Agriculture, Morogoro, Tanzania, fwdulle@sua.ac.tz, orcid.org/0000-0003-2016-6596

Wulystan Pius Mtega

Department of Informatics & Information Technology, College of Natural and Applied Sciences, Sokoine University of Agriculture, Morogoro, Tanzania, wmtega@sua.ac.tz, orcid.org/0000-0001-8471-1878

Abstract

The goal of this research was to investigate the communication channels that enhance data sharing among agriculture researchers in Tanzania. Specifically, the study aimed to identify communication channels that are used by agricultural researchers in Tanzania, examine the extent to which such channels were used to enhance data sharing among agricultural researchers, and examine the factors that influence the choices of channels used in data sharing. A descriptive cross-sectional design, alongside

Communication Channels and their Potential Applicability

quantitative, and qualitative approaches, was employed to collect data from 204 respondents. The Concentric Layered Model for the channel choices was used to guide this study. Results indicate that both mediated, and non-mediated channels existed and were used as data-sharing avenues and channels. The majority of the researchers (77.9%) preferred to use non-mediated channels. Also, the findings indicate that more than 50% of respondents agreed that timely delivery, the cost of the channel, and convenience of a channel were among the factors influencing researchers in their channel selection. It can be concluded therefore that research institutions should invest in mediated channels that have been underutilized to strengthen data-sharing practices among researchers.

Key Words: Agricultural Research; Agricultural Research Data; Communication Channels; Data Sharing; Agricultural researchers

1. Introduction

Research data is the key to scientific agricultural research. Data use in research drives the discovery of new areas of research (Tenopir et al., 2018). Furthermore, data is important for future research and reproducibility analyses (Harper et al., 2018). Access to and use of agricultural research data can drive the development of agricultural science and technology (Zhao & Wang, 2015). Such data should therefore be shared through appropriate communication channels, which serve as pipelines through which messages are transported (Suthers, 2017). In this case, a communication channel links the sender and receiver of data.

Communication channels may be mediated or non-mediated. Mediated communication channels are channels that allow communication between individuals to occur via computer networks and other electronic media (Thompson, 2020; van Deventer & Snyman, 2004; Yao & Ling, 2020). For example, data sharing through mediated communication channels may be through publishing articles on electronic platforms such as AGRA, AGROVOC, and AIMS (Delserone and Dinkelman, 2016; Dodgson et al., 2021). By comparison, non-mediated communication channels include the use of interpersonal communication (Zhao & Wang, 2015). Examples of non-mediated channels are written correspondence and face-to-face dialogue (meetings, discussions). The use of appropriate communication channels guarantees that the data reaches the receiving audience.

The choice of appropriate communication channels is essential in data sharing. A better choice of suitable communication channels ensures the achievement of effective communication (Mtega, 2021). Some criteria are used to guide individuals in the choice of an appropriate channel. A good communication channel does not distort a message and should be fast, confidential, easy to use, accessible, and reachable by both the sender and receiver (Kok et al., 2016; Zizka, 2014).

Over the years, governments and organizations worldwide have made deliberate efforts to embrace data sharing. Funding agencies in the United States have developed policies to encourage research data sharing (Bates, 2017). For example, data sharing has become a criterion for supporting research projects (Bezuidenhout & Chakauya, 2018). Furthermore, the government of Tanzania has made several initiatives to foster data accessibility. The establishment of the Database on Food and Agriculture (Country STAT) aimed at ensuring timely, reliable, and affordable data (Dunmore & Di Cori, 2011; National Bureau of Statistics, n.d.). Moreover, the government of Tanzania implemented the Policy, Regulatory and Institutional Framework to ensure improved data quality and timeliness of access (United Republic of Tanzania, 2015). Likewise, the establishment of the National ICT (Information and Communication Technology) Policy, which focused on areas of ICT infrastructure and universal access to data (<https://www.nbs.go.tz/index.php/en/>), was among the efforts to make data accessible to the public. The establishment of an open data policy draft and an open data portal is among the initiatives intended to make data accessible to the public (Agrawal, 2017). But, until now, the efforts to implement open data in Tanzania has not achieved much success, although the remnants still exist. Until now, different types of data, including agricultural administrative data, and statistical data, are said to be accessible through web-based platforms, websites, and email in the National Bureau of Statistics (NBS) (Bhatia et al., 2016; United Republic of Tanzania, 2015). In addition, weather and climate data that are useful for agricultural research are shared and accessed by the Tanzania Meteorological Agency (TMA) via web-based platforms (Tanzania Meteorological Authority, 2020). Despite these initiatives in Tanzania, the channels for sharing agriculture research data among researchers from research institutions in Tanzania are unknown. There is, therefore, a need to find the channels used for sharing existing data from research institutions.

1.1. Statement of the Problem

Agricultural research is a data-intensive activity. Agricultural researchers use various channels to share their data. In Tanzania, much research has been conducted on communication channels used to share information, and knowledge with researchers, smallholder farmers, and other agricultural stakeholders (Barakabitze et al., 2015; Mtega, 2021; Mtega & Ngoepe, 2019; Ndimbwa et al., 2019, 2021). It can be noted, however, that most of these studies have focused on channels for sharing agricultural information, and knowledge, while none of them focuses on channels for sharing agricultural research data in particular, despite the large amount of research data that exists in agricultural research institutions (Tanzania Agriculture Research Institute, 2019). While there is consequently much knowledge on the effectiveness of channels of various forms in sharing agricultural information and knowledge, there is undoubtedly a lot of untapped potential in terms of how these channels can be equally used to enhance agricultural research data sharing among agricultural researchers. This study was, therefore, set up to fill this knowledge gap by investigating the existing channels in agricultural institutions, their potential applicability in enhancing agricultural research data sharing, and their extent of usage to gauge whether they may be effective and efficient in sharing agricultural research data, and factors that influence researchers in channel choice in Tanzania.

1.2. General Objective

The main goal of this research was to investigate the communication channels that enhance data sharing among agriculture researchers in Tanzania.

1.2.1. Specific Objectives

Specifically the study aimed to:

1. Determine the communication channels that are used by agricultural researchers in Tanzania.
2. Examine the extent to which such channels are used to enhance data sharing among agricultural researchers.
3. Examine the factors that influence the choices of channels used in data sharing.

1.2.2. Research Questions

1. Which communication channels are used by agricultural researchers in Tanzania?
2. To what extent are the channels used to enhance data sharing among agricultural researchers?
3. Which factors influence the choices of channels used in data sharing?

2. Literature Review

The literature review for this study covers themes that are in line with the objectives. It focuses on previous studies that dealt with channels used for sharing agricultural research data, the extent of the channel usage, and the criteria guiding channel selection. This rigorous review contextualized the current study within the wider area of the channel for sharing agricultural research data.

2.1. Communication Channels Used for Sharing Agriculture Research Data

Data is defined as a visual representation of something. Data encompasses observation or information that is obtained from a research project, including experimental samples, technologies, materials, products, and procedures (Kirub, 2016). In this current study, data can be raw or analyzed, quantitative or qualitative, and is collected by researchers and can be shared with other researchers for use or reuse. Research data are data generated through research activities; such data can be in the form of textual, quantitative, and qualitative (Tripathi et al., 2017), and may be in the form of images, recordings, musical compositions, verbal communications, simulations, experimental readings, and code. In agriculture, research data contain crop breeding data, baseline survey data, genomic data, remote sensing data, and geographical information systems (GIS) data (Kirub, 2016). Agricultural research data can be obtained through electronic objects and tools such as sensors, laboratory tools, voice interviews, and online surveys (Boté & Termens, 2019; Zhao & Wang, 2015). Such important data are required to be shared through appropriate channels simply because they are related to agricultural scientific research. Data sharing in this case is defined as a process of

Communication Channels and their Potential Applicability

making data available to others, although not necessarily by depositing it in repositories (Thoegersen & Borlund, 2021). Moreover, data sharing can take place in three methods; data is deposited in a repository, sent upon special request, and published as supplementary materials on the journal website (Gray et al., 2018; Williams, 2022). Likewise, data can be shared in the form of manuscripts, including pre-print products generated by researchers (Jeng et al., 2016).

Communication channels are media that facilitate the transfer of messages from sender to receiver (Mtega, 2021). In this case, therefore, communication channels are also data-sharing channels. However, not all channels work in all environments, and the proper selection of the channel is therefore important to attain effective data delivery. Previous studies have examined different communication channels used for sharing scientific data. In a study that investigated how communication channels are used to increase awareness among researchers, Zaira (2012) observes that the online channels preferred by researchers included blogs, Wiki, Facebook, LinkedIn, and Twitter. The study also reported that the offline channels preferred by researchers in creating awareness among them include printed materials, magazines, and frequent departmental meetings. Researchers from developed nations, in particular, have been sharing and accessing research data through channels that are web-based, such as data repositories. Several meta-analysis and survey studies have revealed that researchers have been sharing research data by depositing it into data repositories (Cooper et al., 2019; Feder et al., 2015; Drakos et al., 2015; Suhr et al., 2020; Thoegersen & Borlund, 2021).

When the data file is shared in repositories, they can be preserved for future use and reuse. The presence of metadata standards and agreements, among other things, facilitated researchers and research institutions in sharing and accessing data from data repositories. In line with this, through reviewing selected publications by Crop Sciences faculty, it was found that data sharing takes place in three methods, namely depositing data into repositories, sending data to researchers upon request, and supplementary materials on journal websites (Williams, 2022). Findings indicate that data sharing in articles appeared in 50 different journals. However, studies from low-income countries reveal that data sharing is not through data repositories; a majority share data through email on request, within published papers, collaborative initiatives, and through research unit servers (Bull et al., 2015;

Katabalwa et al., 2021; Koopman & de Jager, 2016). It is clear that the presence of a data-sharing channel is important to enhance data to be accessible for use by others. Most of the existing studies from developed nations have revealed the presence of formal data sharing channels such as data repositories as compared to most developing countries, where there are no existing data repositories to facilitate data sharing. This therefore called for a detailed study to investigate the current data-sharing channels existing and used by agricultural researchers in less developed countries like Tanzania.

2.2. Extent of Channel Usage and Criteria that Determine Choice of Channels

The selection of proper channels can ensure the message, or in this case, the data reaches the intended audience. The criteria to guide channel selection was studied in detail. In the study that focused on literature review of 36 papers to identify methodologies and knowledge gaps in channel choices by Madsen and Kraemmergaard (2015), the results revealed that the criteria that guide researchers on choices of communication channels include: channel characteristics (interactions, ease of use, usefulness), task characteristics (type of task, adequacy of information), personnel characteristics including social demographics (age, gender, race, education, and income), situational constraints (channel availability, price, and distance to channel) and satisfaction with services offered by the channel. All of these were said to motivate individuals in making their channel selections.

Another study that indicates factors influencing choices of channels is the one by Fichman and Hara (2007), the findings show that channel choice is determined by six criteria: the task at hand, physical proximity, social proximity, sender and receiver accessibility to use the channel, individual preferences about the channel, and the initial channel used. Furthermore, the study that examined the knowledge sharing and channel choice in *The New Way of Working (NWOW)* by Kok et al. (2016), has reported that employees' choices of channels for knowledge sharing were based on factors such as usability, experience with the channel, effectiveness, and convenience. For example, some employees preferred to use multiple channels when sharing knowledge of sensitive information, while some other employees were still working in a more formal and traditional non-*New Way of Working (NWOW)* manner, using fewer channels such as face-to-face communication.

Communication Channels and their Potential Applicability

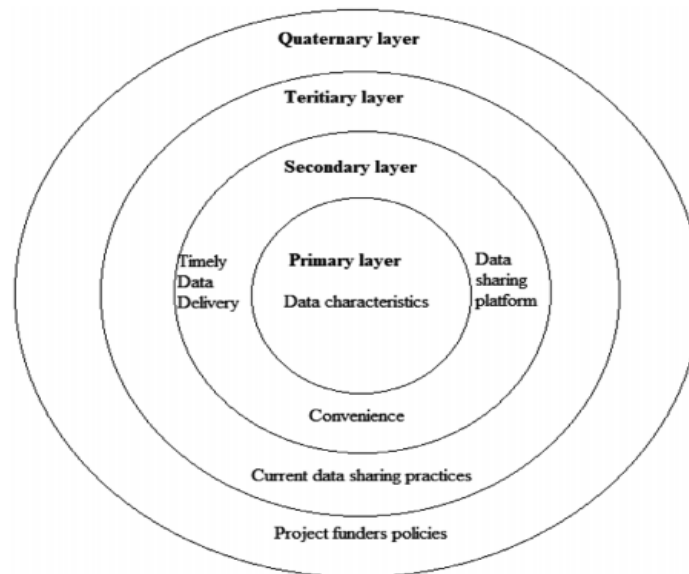
Most published literature on data-sharing channels focuses on research in high-income settings. Most existing studies have revealed that researchers are motivated to deposit their data due to the existing formal data repositories with well-defined metadata standards, file formats, policies, and other agreed arrangements (Cooper et al., 2019; Elsayed, 2018; Federer et al., 2015; Jeng et al., 2016; Thoegersen & Borhund, 2021; Zhao & Wang, 2015). However, although the literature is limited, literature concerning data-sharing channels from low and middle-income settings has exposed several aspects of data sharing (Bull et al., 2015; Katabalwa et al., 2021; Koopman & de Jager, 2016). Earlier literature reveals that the best data-sharing channels that have been operating include the sending of data upon request to researchers who trust each other and work collaboratively, publishing in the papers, and depositing data in repositories. It is interesting to see how Tanzania fits into the general landscape of data sharing in the contexts other developing countries.

2.3. Theoretical Framework

This study adopted the Concentric Layered Model. This model in prior studies was used to explain the factors influencing employees in channel choices when sharing knowledge in work teams and in a *New Way of Working (NWOW)* (Kok et al., 2016; Snyder & Lee-Partridge, 2013). The previous four-level model comprised four concentric circles, each consisting of a group of factors that influence the choice of a particular channel. The first or inner layer (primary layer) includes the type of information, sender efficacy level, and preferences. The secondary level includes richness, ability to keep records, reliability, conveyance of the message, feedback, trust, and preference. The third (tertiary) layer includes embedded practices, the degree of translation and transformation, and the key boundaries. The fourth (quaternary) layer includes the institutional culture, policies, and legal requirements (Snyder & Lee-Partridge, 2013).

In this study, the Concentric Layered Model was modified to take into account the factors influencing the choice of channels in the tertiary and quaternary layers. Based on the modified model (Figure 1 below), the factors that influence researchers' choice of channel for sharing data include the primary layer (data characteristics), secondary layer (timely delivery of data, convenience, and availability of data sharing platforms), tertiary layer – the current

Fig. 1: Adopted and modified model for data sharing channel selection (from Snyder & Lee-Partridge, 2013).



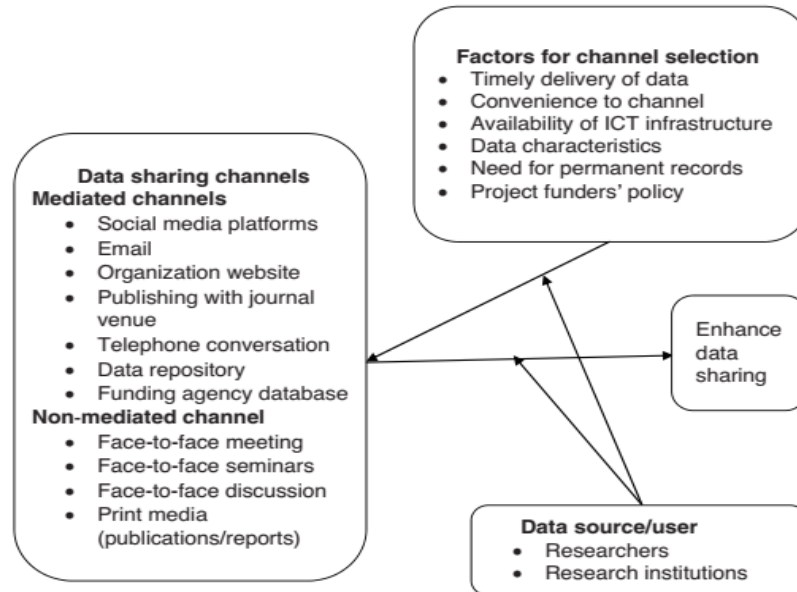
data sharing practice (submission of the data file to institutions or funders for permanent records) and quaternary layer (funding agency policies). Based on Figure 1, the current study findings indicate that some factors concur with the factors in the previous model and the model, therefore, explaining very well the factors influencing the choice of data-sharing channel and thus the model fits for this study.

2.4. Conceptual Framework

The conceptual framework in Figure 2 was developed to explain the linkages between independent, intervening, and dependent variables. The intervening variables include researchers or research institutions that generate, and use research data. When researchers want to share their data, they are guided by some factors in the selection of the channels (independent variables). Data sharing will take place when data files are shared through the selected channels.

Communication Channels and their Potential Applicability

Fig. 2: Conceptual framework showing the relationship exist between the variables for this study.

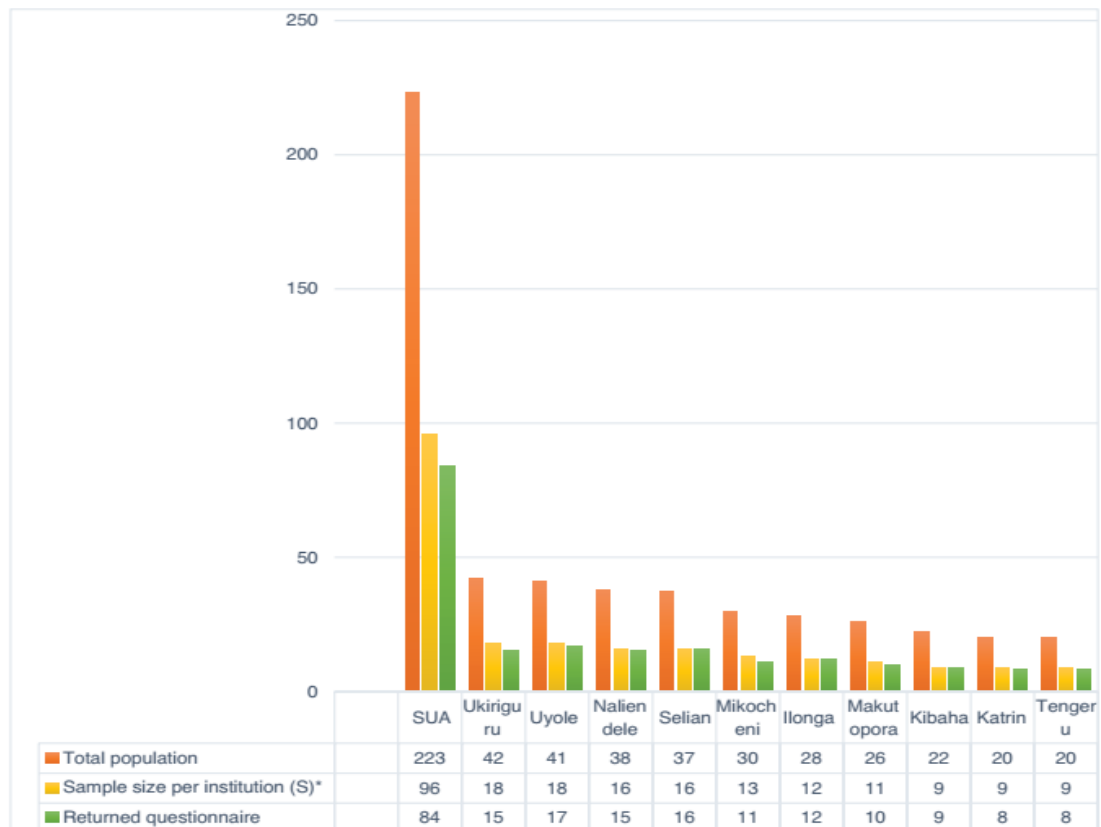


3. Methodology

The study employed a descriptive and cross-sectional research design to investigate communication channels for sharing agricultural research data among agricultural researchers in Tanzania. After receiving research clearance letter from SUA Postgraduate Studies Office, and TARI Headquarter Office the survey was conducted from March 2021 to March 2022. The study employed both qualitative and quantitative approaches in data collection and analysis. The qualitative approach was used to determine opinions, and views about data sharing channels used by researchers, while the quantitative approach was used to examine the existing data sharing and the extent of channel use in data sharing. A combination of the two approaches helped in getting relevant data for the study (Mwinami et al., in press).

The study was conducted in eleven agricultural research institutions in Tanzania. The agricultural research institutions were selected based on the need to have more than twenty (20) researchers at the institution (Figure 3).

Fig. 3: Sampling distribution. Sample size per institution* = study sample size (n) (227) times population per institution over Total population (N).



This number of researchers enabled us to get the intended sample size per institution. The Tanzania Agriculture Research Institution (TARI) centres and Sokoine University of Agriculture (SUA), specifically the Colleges of Agriculture and Veterinary Medicine and Biomedical Sciences, were purposely selected to be included in the study. Key informant interviews were purposively selected to include 10 TARI centre's Directors/Managers and 1 Director of Postgraduate Studies, Research, Technology Transfer & Consultancy for SUA. They were selected because they were thought to be experienced and knowledgeable, and would therefore provide in-depth insights about the topic. The total population of researchers in the study area was 527. The sample size of 227 was obtained using Yamane's (1967) formula for sample size calculation. After obtaining the sample size, a

Communication Channels and their Potential Applicability

representative sample per institution was calculated (Figure 3) (Mwinami et al., in press).

$$n = \frac{N}{1 + N(e)^2}$$

where:

n = Sample size (227)

N = Total study population (527)

e = Level of significance (5% or 0.05)

The data collection used multiple data-gathering techniques to investigate the channels used in sharing data among the researchers. The cross-sectional survey method (questionnaire and interview) and focus group discussions (FGDs) were employed in the study. The questionnaire with closed-ended questions was distributed to 227 agricultural researchers from SUA and ten selected TARI centres. A total number of 204 questionnaires from SUA and ten selected TARI centres were correctly filled, returned, and used in this study. By comparison, qualitative data were collected through interviews with key informants each of which lasted between 15 to 20 minutes (Mwinami et al., in press). A total of two focus group discussions, with each group containing six agricultural researchers, were conducted at Ilonga and Mikochei TARI centres. These research institutions were selected randomly from among several agricultural research institutions. The main purpose was to understand the types of channels used in sharing agricultural research data. During the discussion, respondents were free to express their views, and their opinions were recorded by the researcher. In addition, a document review was made: some of the documents reviewed include the SUA Research Guidelines and Regulations of 2019, the TARI communication strategy of 2019, and the TARI Act of 2016 (Mwinami et al., in press).

3.1. Data Collection and Analysis

Both qualitative and quantitative data were collected and analysed. The present study used content analysis to analyse qualitative data obtained from open-ended questions in the questionnaire, key informant interviews,

and FGDs. The quantitative analysis was accomplished by using Statistical Package for the Social Sciences (SPSS) version 22. In the present study, the descriptive variables for the channels and factors were summarized using a table with the frequencies and percentages. Regarding ethical considerations, the researcher avoided any risk of harm during data collection and ensured the right to privacy and informed consent before engaging any respondent in the study (Mwinami et al., in press).

4. Results and Discussion

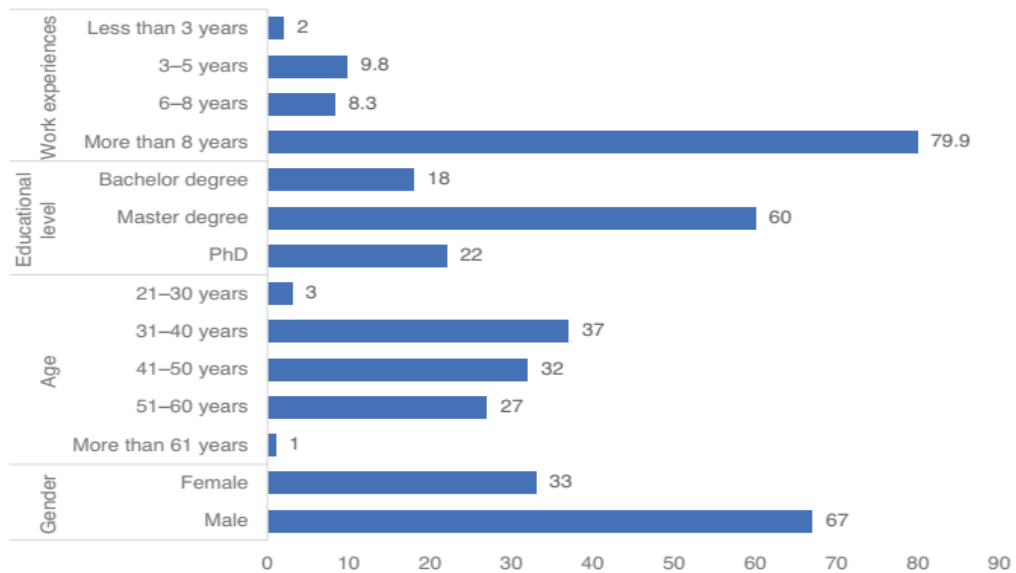
This section presents the findings based on study objectives, channels for data sharing, and the factors influencing agricultural researchers' choices of data sharing channels.

4.1. Demographic Characteristics of Respondents

The demographic characteristics of respondents are presented in Figure 4. Results indicate that a significant percentage (67%) of the researchers who participated in this study were male. This implies that the majority of researchers in the surveyed research institutions were male. This trend suggests that females were in a small proportion (33%) of those employed and therefore participating in agricultural research in the surveyed institutions. The findings further show that most (69%) of the respondents were aged between 31 to 50; the remainder (27%) was aged 50 and above. This implies that significant proportions of researchers were at an active age and engaged in agricultural research. Furthermore, the results also indicate that a large percentage of the respondent (82%) are researchers who had already received their Master's degree and above, while a few (18%) had only completed a Bachelor's degree. This implies that most of the researchers in the surveyed area had a higher level of education that enabled them to conduct research. The results further indicate that a large percentage of researchers (95.3%) had work experience of more than (6) years, while only 10% had work experience of between 1 and 5 years. This implies that the majority of the researchers had a reasonable amount of work experience in conducting research activities and the junior researchers would be gaining experience from seniors (Mwinami et al., in press).

Communication Channels and their Potential Applicability

Fig. 4: Characteristics of the respondents (n=204).



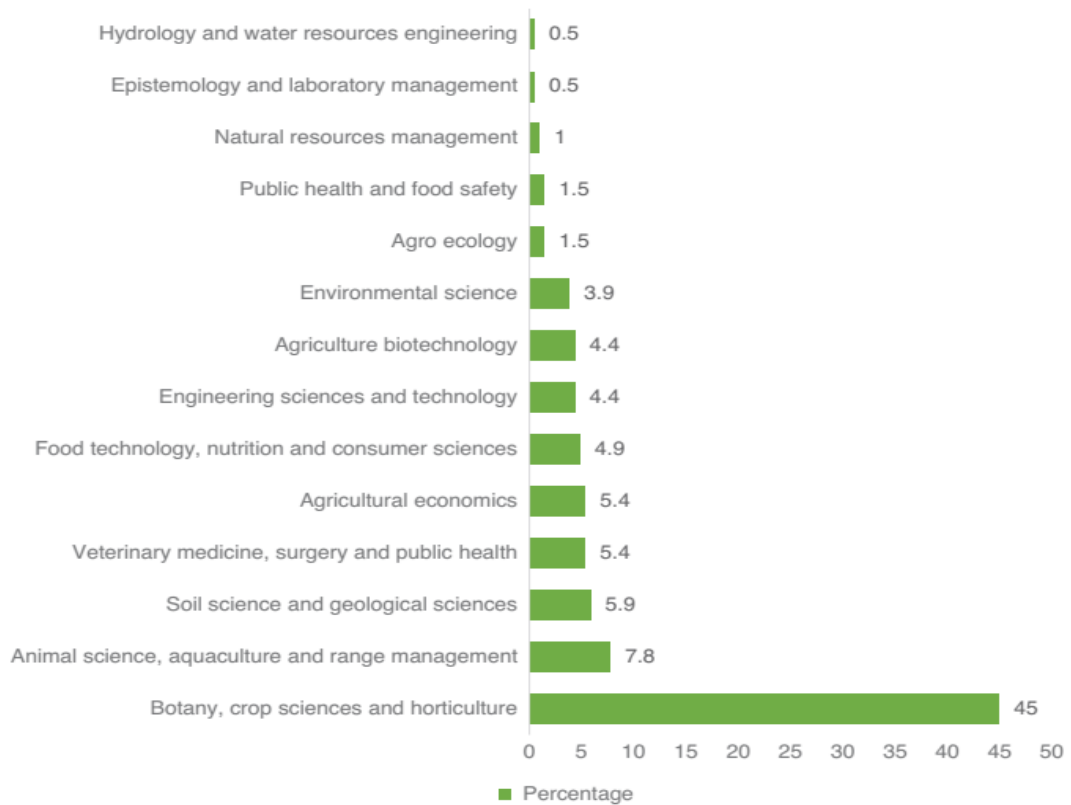
4.2. Researchers' Academic Field of Specialization

The findings in Figure 5 indicate the researchers' academic field of specialization. From the findings, the largest group (45%) of researchers specialized in botany, crop sciences, and horticulture. An additional 7.8% of the respondents specialized in Animal Science, Aquaculture, and Range Management. This implies that researchers who participated in this study specialized in various research fields (Figure 5).

4.3. Channels Existing and Used in Sharing Agricultural Research Data

These findings in Table 1 below reveal that respondents made the heaviest use (77.9% and above) of non-mediated channels, including face-to-face dialogue (meetings, discussions, and seminars) and print media, such as technical reports, etc. These results imply that researchers preferred to use channels that give the chance for face-to-face dialogue, and interaction, and provide timely feedback among researchers. In addition to that, the study established

Fig. 5: Researchers' Academic Field of Specialization (n=204).



that researchers mostly preferred channels that could deliver data promptly. These findings, concurring with the study by Shen (2017), show clearly that interpersonal discussion was used to help in providing data context, declaring assumptions, clarifying jargon, elaborating data structures, and identifying points of interest for effective data sharing. Similarly, findings are related to previous studies which reveal that face-to-face meetings and conversations are the best for research data sharing among researchers (Howland et al., 2015; Shen, 2017). It has also been reported that researchers in universities and research institutions have used formal seminars, workshops, mobile phones, and open discussions as the most preferred and available channels to exchange knowledge and ideas with colleagues (Delserone & Dinkelman, 2016; Kigatiira et al., 2018).

Communication Channels and their Potential Applicability

Table 1: Channels existing and used in agriculture research institutions (n=204).

Data sharing channel	Frequency	Percentage
Mediated channels		
Email	152	74.5
Publishing in online journal platforms	148	72.5
Organization website	130	63.7
Telephone (interpersonal mediated conversation)	107	52.5
Social media platforms	105	51.5
Funding agency database/repositories	84	41.4
Institutional repository	74	36.3
Non-mediated channels		
Face-to-face meeting	182	89.2
Print media publication (technical reports)	178	87.3
Face-to-face seminars	169	82.8
Face-to-face discussions	159	77.9

The findings in Table 1 show that respondents indicated the strongest preferences for mediated channels such as email, and publication on online journal platforms (74.5% and 72.5%, respectively). These findings imply that preference for such channels is a result of the timely delivery of data due to the available data in electronic format. Researchers had the habit of sharing their data files with fellow researchers through their emails. Researchers' preference to share data files through online journal platforms was motivated by the opportunity they got to increase their data visibility. Moreover, in agriculture research institutions/universities (in this case, at SUA), publishing articles with data in the journal avenue allows for advancement and promotions in the academic ranks. These findings are in line with a previous study by Zhang et al. (2016), which informed that emails were among the best channels used by researchers in sharing data. Concerning online journal publications, other previous studies have also reported the use of this channel (Shen, 2017; Tedersoo et al., 2021).

The study results show that some mediated channels were least used by researchers. These included websites, social media platforms, institutional repositories, and funding agency databases. This underutilization of mediated data sharing channels might have been caused by the absence of data

repositories among research institutions, and a lack of awareness among researchers of using such channels. Most of the research institutions in Tanzania do not have well-developed ICT infrastructures, especially data repositories; this might have hindered researchers sharing their agriculture research data through data repositories.

Findings in Table 1, showing multiple responses, indicate that both mediated and non-mediated channels were used in data sharing at varying levels. In support of this argument, one of the key interviewees at the TARI Selian centre commented:

“I share my research data through the print media (for example technical reports) by submitting it to TARI headquarters. I have usually attempted to share my data file with my fellow researchers through our WhatsApp group. The WhatsApp group enables us to share any fact or data files based on challenges on emerging issues in crops under mandate such as wheat, barley, and maize etc.”

Another key interviewee at TARI Katrin centre commented that:

“The project funders, collaborators and partners usually establish electronic platforms for communication throughout the whole period of the project. You are given a password and username for logging into the platform. Throughout the project, you will share data, information, and knowledge via this platform. At the end of the project, sometimes you are required to submit all the data files via this platform and sometimes data analysis is done by the project funder.”

Furthermore, another key interviewee on channels for data sharing at SUA commented that:

“Currently, in Tanzania, there are no generalist or domain data repositories or platforms maintained by agricultural research institutions. The SUA research regulation and guidelines spell out that data and other research materials generated by researchers belong to the institutions. If one needs to share data at individual, national and international levels are required to follow institutional and national guidelines. There are observed formal and informal data sharing within department and institutional levels. Some data embedded in research outputs are shared

Communication Channels and their Potential Applicability

by depositing through SUA institutional repository as channels that host content from different agriculture research fields. However, agricultural researchers share and access their data from other local and international research institutions, research partners, research collaborators, and agriculture-affiliated agencies. Also, some agriculture researchers access and share their agriculture data through generalist databases hosted by the National Bureau of Statistics (NBS). Another domain database that stores climatic data is hosted by the Tanzania Meteorological Agency (TMA)."

In the same context, through focus group discussions held at Ilonga and Mikocheni, respondents reported the following:

"After the discussion in all the two sessions, researchers revealed the channel they use in sharing their research data, including field days at research institutes. The use of technical reports that contain some agriculture research data in the form of data files are shared via emails and in hard copies also submitted to the TARI headquarters. The annual meetings held in the zonal selected research institutions have been useful in sharing raw data, processed data, and other technologies innovated on the basis of crop specialization, for example maize, grain legumes, sunflower, sorghum and millets."

The second objective of this study was to determine the extent of channel usage to gauge whether they may be effective and efficient in sharing agricultural research data. Through cross-tabulation, the usage of each channel is shown for each of the research institutions. The findings in Table 2 below show that SUA had reported the use of all the existing data-sharing channels, and the extent of usage is higher than at the other research institutions investigated. Of all the data sharing channels that exist within the studied institutions, meetings, conferences, and the organizations' accessible website had a high statistical significance difference at a 1% level of significance ($p=0.005$ and 0.001 , respectively). Moreover, workshops, telephone-mediated conversations, and print media such as institutional publications, including technical reports that carry data had a statistically significant difference at a 5% level of significance ($p=0.049$, 0.045 , 0.045 , and 0.025 respectively). These findings imply that these channels or avenues were preferred over other channels in the studied area. Agricultural researchers at SUA and TARI centres always work in teams; this explains the role of meetings, conferences, telephone conversations, project reports, and workshops for data sharing. Findings show

Table 2: *Extent of channel usage across research institutions as segregated by institution (n = 204).*

Data sharing channel	SUA		Ilonga		Naliendele		Selian		Uyole		Tengeru		Katrin		Ukinguru		Maktopora		Mikochehi		Kibaha		P-value
	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%	
Face-to-face conversation	48	35.6	8	5.9	7	5.2	12	8.9	13	9.6	8	5.9	5	3.7	13	9.6	8	5.9	7	5.2	6	4.4	0.066
Discussions	48	39	6	4.9	13	10.6	8	6.5	8	6.5	6	4.9	4	3.3	10	8.1	5	4.1	10	8.1	5	4.1	0.358
Meetings and conferences	52	35.9	11	7.6	15	10.3	12	8.3	8	5.5	7	4.8	3	2.1	14	9.7	7	4.8	10	6.9	6	4.1	0.005**
Seminars	49	38.3	8	6.2	6	4.7	12	9.4	9	7	7	5.5	3	2.3	11	8.6	9	7	8	6.2	6	4.7	0.098
Workshops	67	36.6	11	6	16	8.7	14	7.7	16	8.7	8	4.4	8	4.4	13	7.1	10	5.5	11	6	9	4.9	0.049*
Email	45	37.5	8	6.7	7	5.8	7	5.8	7	5.8	5	4.2	5	4.2	12	10	9	7.5	10	8.3	5	4.2	0.008
Telephone conversations	44	41.5	8	7.5	6	5.7	5	4.7	12	11.3	2	1.9	4	3.8	7	6.6	8	7.5	8	7.5	2	1.9	0.045*
Social media platforms	43	43	7	7	8	8	5	5	5	5	3	3	4	4	8	8	6	6	4	4	7	7	0.555
Organization website	38	32.5	5	4.3	12	10.3	10	8.5	13	11.1	5	4.3	2	1.7	15	12.8	8	6.8	3	2.6	6	5.1	0.001**
Institutional repository	38	52.8	0	0	5	6.9	3	4.2	5	6.9	4	5.6	3	4.2	6	8.3	2	2.8	4	5.6	2	2.8	0.154
Funding agency database	23	41.8	3	5.5	5	9.1	3	5.5	3	5.5	0	0	1	1.8	6	10.9	3	5.5	4	7.3	4	7.3	0.619
Publishing with journal venues	60	40.5	10	6.8	11	7.4	11	7.4	13	8.8	8	5.4	5	3.4	11	7.4	6	4.1	7	4.7	6	4.1	0.789
Institutional publications	63	37.7	10	6	16	6.9	11	6.6	16	9.6	8	4.8	7	4.2	9	5.4	9	5.4	10	6	8	4.8	0.045*
Project technical reports	61	35.3	10	5.8	16	9.2	14	8.1	16	9.2	8	4.6	8	4.6	14	8.1	8	4.6	10	5.8	8	4.6	0.024*

Note: *Significant at 5% level, **Significant at 1% level, F=Frequency and %=Percentage.

Communication Channels and their Potential Applicability

that there were variations in the existence of channels and the extent to which such channels were used, with SUA showing the primacy in all the channels used for sharing agricultural research data. This might be due to SUA having better ICT infrastructure compared to the other research institutions. SUA also has a large number of researchers from different research fields that would use a multiplicity of channels in sharing data files.

4.4. Factors Influencing Researchers in their Channel Selection

The third objective of this study was to examine factors influencing the choices of data-sharing channels as reflected in the Concentric Layered Model in Figure 1. Findings in Table 3 indicate that more than 50% of respondents shown strongly agree that timely delivery, the cost of the channel, and the convenience of the channel were among the factors influencing researchers' choice of channels for data sharing. Timely delivery of data is among the motivating factors for researchers to use avenues such as face-to-face dialogue that provide timely delivery of data. For example, researchers preferred to use face-to-face dialogue (meetings, seminars, and discussions) because they were among the avenues which provide a chance for the timely exchange of data files, technologies, and innovations. With the advancement of technology, the weather forecast has been improved through the use of the mobile phone in data sharing. The use of CAMALIOT Android apps that uses a cell phone to support scientific research has been used in weather data collection

Table 3: Factors that influenced researchers in channel selection (n=204).

Factors that influence channel selection	Strongly Agree		Agree		Undecided		Disagree		Strongly Disagree	
	F	%	F	%	F	%	F	%	F	%
Timely delivery of data	154	75.5	-	-	36	17.6	14	6.9	-	-
Convenience to channel	101	49.5	77	37.7	20	9.8	6	2.9	-	-
ICT infrastructure	93	46.6	59	28.9	28	13.7	20	8.5	4	1.7
Need for permanent records	75	36.8	80	39.2	40	19.6	9	4.4	-	-
Funders policies	66	32.4	98	48	32	15.7	7	3.4	1	0.5
Data characteristics	43	21.1	94	46.1	39	19.1	27	13.2	1	0.5

and exchange. The app facilitates citizen and researchers to access and collect raw GPS satellite data from individuals with cell phones thus a cell phone act as a channel for transferring data to servers for permanent preservation for future use (See et al., 2022).

The convenience of the channel was observed to encourage researchers to choose a channel. The meetings convened in agricultural research institutions are convenient for researchers. Regular technical meetings for project evaluation (seminars and discussions held in research institutions) are convenient to most researchers to the extent that they motivate researchers in sharing data related to technologies or innovations. Likewise, almost every researcher owns a cell phone which makes it easy to call fellow researchers participating on the same project team or collaborators who are found in another institution just for sharing experiences in particular research projects. It is during such conversations that research data files are shared and hence exchanged using such channels. Elia (2018) recommended the use of mobile phone devices in meteorological stations to facilitate the timely collection and recording of weather raw data. The phone is a channel that can help to make data to be timely shared. In line with this, a study by Msemu et al. (2021) reveals that cell phone has been very useful to disseminate and receive weather data and warning information from weather stations located in different regions in Tanzania.

Data is characterised by its need for permanent records, availability of ICT infrastructures, and project funders' policies; yet these are also among the factors indicating less than 50% slightly in motivating researchers in their choice of channel. The availability of ICT infrastructure such as databases, and repositories that are connected by the internet motivates researchers to share their data by depositing their data files in funders' databases at the end of the project. Other data files have been published as supplementary materials in journal e-platforms. The need for permanent records for data obtained from research activities is among the factors that motivate researchers to choose a channel or avenue to share data. For example, by sending their data files to the funder's databases and research institutions, the researchers' main intention is to have permanent records for the data obtained during research projects. Therefore, the possession of data repositories or databases by the project funder motivates researchers to share their data through the available avenue. These findings are in line with the prior studies which have revealed that the availability of funding agencies' databases was observed to facilitate

Communication Channels and their Potential Applicability

researchers to select funder's databases as their channel to share their data (Bezuidenhout & Chakauya, 2018; Dodgson et al., 2021; Suhr et al., 2020).

5. Conclusion

Access to agricultural research data is essential for effective agricultural research. Our study findings reveal that agricultural researchers use both mediated and non-mediated channels for sharing their data. The non-mediated channels, such as meetings, seminars, and print media (institutional publications or technical reports), were mostly preferred by researchers. This implies that data sharing in research institutions is still happening via direct contact and face-to-face dialogue rather than via electronic platforms. Our findings indicate that timely delivery, the cost of the channel, and the convenience of the channel were among the factors influencing researchers to choose the channels for data sharing. From the findings, it is implied that data sharing would improve if research institutions could improve the use of digital technological channels in data sharing. This study contributes to the body of knowledge through the proposed conceptual framework for data-sharing channels because it can be used to explain existing data-sharing channels and the criteria for their selection.

6. Recommendations

Based on the study objectives and conclusion, the following are our recommendations.

1. The study recommends that, research institutions should devote themselves to preparing adequate funds for investing in mediated data-sharing channels such as generalised and specialised agriculture data repositories. Such investments could strengthen data practice sharing among researchers at local and international levels.
2. Research institutions should strengthen the non-mediated data-sharing channels that are currently used by researchers so that they continue to facilitate data sharing among researchers. The study recommends that research institutions should recognise the role of libraries in facilitating data management and sharing. A library is a

place where some of the mediated and non-mediated data-sharing channels are hosted. The data repositories and other print research outputs and other resources embedded with research data can be stored in a library. Therefore, strengthening the existing library infrastructures and building new libraries in agricultural research institutions can have a greater impact on data management sharing and use/reuse practices. The improvement of libraries should go hand in hand with recruiting librarians and ICT technicians who would play a greater role in data management and facilitate data sharing among researchers.

3. Research institutions should work closely with other international research institutions or partnerships so that researchers are given the opportunity to access and share their data with fellow researchers through international web-based channels.
4. Research guidelines, and regulations in agricultural research institutions should incorporate, and declare the formal channels for sharing research data.
5. This study further recommends that a data sharing policy framework should be amended to guide both the formal and informal data sharing procedures, and state clearly proper use of the mediated and non-mediated channels in data sharing among agricultural researchers.

Acknowledgment

Sincere gratitude goes to my study supervisors from the Sokoine University of Agriculture, and to my employer Jordan University College for allowing me to pursue further studies at the Sokoine University of Agriculture (SUA).

References

Agrawal, A. (2017). *Data roadmaps for sustainable development: Assessment and lessons learned*. Global Partnership for Sustainable Development Data. https://www.data4sdgs.org/sites/default/files/services_files/RoadmapsAssessmentReport_Dec2017_FINAL.pdf

Communication Channels and their Potential Applicability

Barakabitze, A. A., Kitindi, E. J., Sanga, C., Shabani, A., Philopo, J., & Kibirige G. (2015). New technologies for disseminating and communicating agriculture knowledge and information: Challenges for Agricultural Research Institutes in Tanzania. *The Electronic Journal of Information Systems in Developing Countries*, 70(2), 1–22. <https://doi.org/10.1002/j.1681-4835.2015.tb00502.x>

Bates, J. (2018). The politics of data friction. *Journal of Documentation*, 74(2), 412–429. <https://doi.org/10.1108/JD-05-2017-0080>

Bezuidenhout, L., & Chakauya, E. (2018). Hidden concerns of sharing research data by low/middle-income country scientists. *Global Bioethics*, 29(1), 39–54. <https://doi.org/10.1080/11287462.2018.1441780>

Bhatia, V., Stout, S., Baldwin, B., & Homer, D. (2016). *Results data initiatives: Findings from Tanzania*. Development Gateway. <https://developmentgateway.org/wp-content/uploads/2020/10/RDI-Tanzania.pdf>

Boté, J.-J., & Termens, M. (2019). Reusing data technical and ethical challenges. *DESIDOC Journal of Library and Information Technology*, 39(6), 329–337. <https://doi.org/10.14429/djlit.39.06.14807>

Bull, S., Cheah, P. Y., Denny, S., Jao, I., Marsh, V., Merson, L., More, N. S., Thanh Nhan, L. N., Osrin, D., Tangseefa, D., Wassenaar, D., & Parker, M. (2015). Best practices for ethical sharing of individual-level health research data from low- and middle-income settings. *Journal of Empirical Research on Human Research Ethics*, 10(3), 302–313. <https://doi.org/10.1177/1556264615594606>

Cooper, D., & Springer, R. (2019). *Data communities: A new model for supporting STEM data sharing [Issue Brief]*. ITHAKA S+R <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1110&context=scholcom>

Delserone, L. M., & Dinkelman, A. L. (2016). *Investigating the practices and needs of agricultural researchers at the University of Nebraska-Lincoln*. University of Nebraska-Lincoln. <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1360&context=libraryscience>

Dodgson, J. E., Gill, S., Marinelli, K., Duckett, L., Chetwynd, E., Bamberger, E., Garcia, C., Bamberger, C., & Costello, R. (2021). Policy statement a data sharing and data accessibility. *Journal of Human Lactation*, 37(3), 447–448. <https://doi.org/10.1177/08903344211016291>

Drakos, A., Protonotarios, V., & Manouselis, N. (2015). AgINFRA: a research data hub for agriculture, food and the environment. *F1000Research*, 4(127), 1–19. <https://doi.org/10.12688/f1000research.6349.1>

Dunmore, J., & Di Cori, S. (2011). *CountrySTAT for Sub-Saharan Africa*. The United Nations Food and Agriculture Organization. <https://www.fao.org/3/bd063e/bd063e.pdf>

Elia, E. F. (2018). Towards establishing an effective data management system in Tanzania: A comparative analysis of scientific climatic data and farmers' perception

of climate change and variability. *University of Dar es Salaam Library Journal*, 13(1), 36–53.

Elsayed, A. M. (2018). Research data management and sharing among researchers in Arab universities: An exploratory study. *IFLA Journal*, 44(4), 281–299. <https://doi.org/10.1177/0340035218785196>

Federer, L. M., Lu, Y.-L., Joubert, D. J., Welsh, J., & Brandys, B. (2015). Biomedical data sharing and reuse: Attitudes and practices of clinical and scientific research staff. *PLoS One* 10(6), Article e0129506. <https://doi.org/10.1371/journal.pone.0129506>

Fichman, P., & Hara, N. (2007). Behavior complexity theory of media selection: A proposed theory for global virtual teams. *Journal of Information Science*, 33(1), 63–75. <https://doi.org/10.1177/0165551506068145>

Gray, B., Babcock, L., Tobias, L., McCord, M., Herrera, A., Osei, C., & Cadavid, R. (2018). *Digital farmer profiles: Reimagining smallholder agriculture*. Feed the Future, The U.S Government Global Hunger and Food Security Initiative. https://grameenfoundation.org/documents/Data_Driven_Agriculture_Web.pdf

Harper, L., Campbell, J., Cannon, E. K. S., Poelchau, M., & Andorf, C. (2018). AgBioData Consortium recommendations for sustainable genomics and genetics databases for agriculture. *Database*, 1(32), 1–32. <https://doi.org/10.1093/database/bay088>

Howland, F., Muñoz, L. A., Staiger-Rivas, S., Cock, J., & Alvarez, S. (2015). Data sharing and use of ICTs in agriculture: working with small farmer in Colombia. *Knowledge Management for Development Journal*, 11(2), 44–63.

Jeng, W., He, D., & Oh, J. S. (2016). Toward a conceptual framework for data sharing practices in social sciences: A profile approach. *Proceedings of the Association for Information Science and Technology*, 53(1), 1–10. <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pr2.2016.14505301037>

Katabalwa, A. S., Bates, J., & Abbott, P. (2021). Potential opportunities and risks of sharing agricultural research data in Tanzania. *IASSIST Quarterly*, 45(3-4), 1–8. <https://doi.org/10.29173/iq997>

Kigatiira, K. K., Mberia, H. K., & wa Ngula, K. (2018). The effect of communication channels used between extension officers and farmers on the adoption of Irish potato farming. *International Journal of Academic Research in Business and Social Science*, 8(4), 373–387. <https://doi.org/10.6007/ijarbss/v8-i4/4020>

Kirub, A. (2016). *Agricultural research data management: Principles, policy, and practice*. Ethiopian Institute of Agriculture Research. https://www.researchgate.net/publication/308030253_agricultural_research_data_management_principles_policy_and_practice

Kok, A., Bellefroid, B., & Helms, R. W. (2016). Knowledge sharing and channel choice: Effects of the New Way of Working. *ECKM 2013 – Proceedings of the 14th European*

Communication Channels and their Potential Applicability

Conference on Knowledge Management, Lithuania, 2, 849–859. <https://doi.org/10.13140/RG.2.1.1533.2724>

Koopman, M. M., & de Jager, K. (2016). Archiving South African digital research data: How ready are we? *South African Journal of Science*, 112(7–8), Article 2015-0316. <https://doi.org/10.17159/sajs.2016/20150316>

Madsen, C., & Kraemmergaard, P. (2015). Channel choice: A literature review. In E. Tambouris, M. Janssen, H. J. Scholl, M. A. Wimmer, K. Tarabanis, M. Gascó, B. Klievink, I. Lindgren & P. Paycek (Eds.), *Electronic Government. EGOV 2015: Vol. 9248* (pp. 3–18). Springer. https://doi.org/10.1007/978-3-319-22479-4_1

Msemu, H. E., Taylor, A. L., Birch, C. E., Dougill, A. J., & Hartley, A. (2021). The value of weather and climate information to the Tanzanian disaster risk reduction sector using Nonmonetary Approaches. *Weather and Climatic Society*, 13(4), 1055–1068. <https://doi.org/10.1175/WCAS-D-21-0005.1>

Mtega, W. P. (2021). Communication channels for exchanging agricultural information among Tanzanian farmers: A meta-analysis. *International Federation of Library Associations and Institutions*, 47(4), 570–579. <https://doi.org/10.1177/03400352211023837>

Mtega, W. P., & Ngoepe, M. (2019). A framework for strengthening agricultural knowledge systems for improved accessibility of agricultural knowledge in Morogoro Region of Tanzania. *Journal of Librarianship and Information Science*, 51(3), 629–642. <https://doi.org/10.1177/0961000617742456>

Mwinami, N. V., Dulle, F. W., & Mtega, W. P. (2022). Data preservation practices for enhancing agricultural research data usage among agricultural researchers in Tanzania. *Journal of Librarianship and Information Science*. <https://doi.org/10.1177/09610006221138110>

National Bureau of Statistics. (n.d). *Dissemination and Pricing Policy: The NBS recognizes that timely dissemination of data to specialized users for research purposes and other decision makers is beneficial*. Retrieved April 20, 2022, from <https://www.nbs.go.tz/index.php/en/about-us/policies-and-legislations/policies/12-dissemination-and-pricing-policy>

Ndimbwa, T., Mwantimwa, K., & Ndumbaro, F. (2021). Channels used to deliver agricultural information and knowledge to smallholder farmers. *IFLA Journal*, 47(2), 153–167. <https://doi.org/10.1177/0340035220951828>

Ndimbwa, T., Ndumbaro, F., & Mwantimwa, K. (2019). Delivery mechanisms of agricultural information and knowledge to smallholder farmers in Tanzania: A meta-analysis study. *University of Dar es Salaam Library Journal*, 14(2), 87–98.

See, L., Soja, B., Heyl, A., & Mickein, I. (2022, March 17). *Use cellphone to improve weather forecasts*. International Institute for Applied Systems Analysis. <https://iiasa.ac.at/news/mar-2022/use-your-cellphone-to-improve-weather-forecasts>

Shen, Y. (2017). Data sharing practices, information exchange behaviors, and knowledge discovery dynamics: a study of natural resources and environmental

- scientists. *Environmental Systems Research*, 6(9), 1–14. <https://doi.org/10.1186/s40068-017-0086-5>.
- Snyder, J., & Lee-partridge, J. E. (2013). Understanding communication channel choices in team knowledge sharing. *Corporate Communications: An International Journal*, 18(4), 417–431. <https://doi.org/10.1108/CCIJ-03-2012-0026>
- Suhr, B., Dungal, J., & Stocker, A. (2020). Search, reuse and sharing of research data in materials science and engineering: A qualitative interview study. *PLoS One* 15(9), Article e0239216. <https://doi.org/10.1371/journal.pone.0239216>
- Suthers, A. L. (2017). *Evaluating Effective Communication Methods: Improving Internal Communication* [Electronic Theses and Dissertations, East Tennessee State University]. Digital Commons @ East Tennessee State University. <https://dc.etsu.edu/etd/3293/>
- Tanzania Agriculture Research Institute. (2019). *Communication strategy*. <http://www.tari.go.tz/assets/uploads/documents/c2cd7f2933a36faf8939d3ec8616c360.pdf>
- Tanzania Meteorological Authority. (2020). *Tanzania meteorological authority (TMA) Agro meteorological database*. http://agromet.meteo.go.tz/storage/docs/TMA_Agromet_Database_Manual.pdf
- Tedersoo, L., Kungas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, A., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K., & Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data*, 8(192). <https://doi.org/10.1038/s41597-021-00981-0>
- Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Earth and space science research data sharing: practices and attitudes of geophysicists. *Earth and Space Science*, 5(12), 891–902. <https://doi.org/10.1029/2018EA000461>
- Thoegersen, J. L., & Borlund, P. (2021). Researcher attitudes toward data sharing in public data repositories: a meta-evaluation of studies on researcher data sharing. *Journal of Documentation*, 78(7), 1–17. <https://doi.org/10.1108/JD-01-2021-0015>
- Thompson, J. B. (2020). Mediated Interaction in the Digital Age. *Theory, Culture and Society*, 37(1), 3–28. <https://doi.org/10.1177/0263276418808592>
- Tripathi, M., Shukla, A., & Sonker, S. K. (2017). Research data management practices in University libraries: A Study. *DESIDOC Journal of Library & Information Technology*, 37(6), 417–424. <https://doi.org/10.14429/djlit.37.6.11336>
- United Republic of Tanzania. (2015). *Agricultural sector development-II 2015/2016-2024/2025*. <https://faolex.fao.org/docs/pdf/tan160643.pdf>
- van Deventer, J. P., & Snyman, M. M. M. (2004). Computer-mediated communication and knowledge sharing in a South African context. *South African Journal of Information management*, 6(1), Article a292. <https://sajim.co.za/index.php/sajim/article/view/292/283>

Communication Channels and their Potential Applicability

Williams, S. C. (2022). Data practices ten years later: A new review of selected publications by crop sciences faculty. *Journal of Agricultural & Food Information*, 23(1-2), 1-37. <https://doi.org/10.1080/10496505.2021.2013850>

Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper and Row.

Yao, M. Z., & Ling, R. (2020). What is computer-mediated communication? An introduction to the special issue. *Journal of Computer-Mediated Communication*, 25(1), 4-8. <https://doi.org/10.1093/jcmc/zmz027>

Zaira, M. (2012). *Efficient communication channels towards the researchers*. [Unpublished doctoral dissertation] Laurea University of Applied Sciences, Otaniemi. <https://www.theseus.fi/bitstream/handle/10024/46477/Zaira%20Mammadova-%20Thesis.pdf?sequence>

Zhang, Y., Wang, L., & Duan, Y. (2016). Agricultural information dissemination using ICTs: A review and analysis of information dissemination models in China. *Information Processing in Agriculture*, 3(1), 17-29. <https://doi.org/10.1016/j.inpa.2015.11.002>

Zhao, H., & Wang, J. (2015). Development and current situation of agricultural scientific data sharing in China. In D. Li & Y. Chen (Eds.), *IFIP Advances in information and communication technology* (Vol. 452, pp. 80-86). Springer Publishing. https://doi.org/10.1007/978-3-319-19620-6_10.

Zizka, L. (2014). *Communication channels: The effects of frequency, duration and function on gratification obtained* [Doctoral dissertation, Walden University]. Walden University ScholarWorks. <https://scholarworks.waldenu.edu/dissertations/83/>

CHAPTER FIVE

PAPER IV

Factors influencing data sharing practises, perceptions and motivations for improved accessibility of agricultural research data in Tanzania.

Nolasko Victory Mwinami,

PhD candidate, Department of Informatics & Information Technology (DIIT),
College of Natural and Applied Sciences (CoNAS), Sokoine University of Agriculture
(SUA), Morogoro, Tanzania, mwinaminolasko@yahoo.com

Frankwell W. Dulle

Associate professor at the Department of Informatics & Information Technology (DIIT),
College of Natural and Applied Sciences (CoNAS), Sokoine University of Agriculture
(SUA), Morogoro, Tanzania, fwdulle@sua.ac.tz

Wulystan Pius Mtega

Associate professor at the Department of Informatics & Information Technology (DIIT),
College of Natural and Applied Sciences (CoNAS), Sokoine University of Agriculture
(SUA), Morogoro, Tanzania
wmtega@sua.ac.tz

Targeted Journal and status: HSOA Journal of Environmental Science Current Research, 6:045, DOI: 10.24966/ESCR-5020/100045, Paper is accepted for publication.

Abstract

Accurate and reliable information on the data sharing practices and what motivates researchers to share their data is important in making data accessible for use/reuse. The purpose of this paper is to report a study that investigated factors influencing data-sharing practices, perceptions, and motivations for enhanced accessibility of agricultural research data among agricultural researchers in Tanzania. This research employed a cross-sectional survey method to examine data-sharing practices in agricultural research institutions and examined factors motivating researchers in their data-sharing. The sample size included 227 while from this sample size some 204 respondents participated in the study. The findings have revealed that more than 80% of the researchers shared their data with project funders, research institutions, and with departmental colleagues. The findings also indicate that many researchers (60%) shared their data with the public through media, institutional libraries, and online publications. The findings reveal that both individual and institutional factors influence researchers in data sharing. However, some of the factors were statistically significant at 5%; these are perceived community benefit, perceived ability to share, data sharing norms, perceived risk, and funding agency policies. In addition, the findings indicate that agricultural researchers face several challenges in their data sharing, including a lack of data sharing policy that could enhance data sharing in research institutions. This study has practical implications for promoting more reliable, and beneficial data sharing in the agricultural scientific community. The findings of this research can facilitate the establishment of data-sharing policies and data repositories that could improve easy data sharing. The present study has significant theoretical contributions: the TPB model has been used in this study to explain the factors motivating researchers in their data-sharing behaviour.

Keywords

Agricultural Research, Agriculture Research Data, Data Sharing, Research Data.

1.0 Introduction

Agriculture research depends on scientific data. Scientific data are data generated through research activities (Zhao and Wang, 2015). Such data are needed to be shared to attain the benefits accrued. Scientific data sharing contributes to the improvement of scientific research. Thus, data sharing enhances the advances in agricultural research in the end leading to an increase in food production (Harper *et al.*, 2018; Kirub, 2016b; Moore *et al.*, 2021; OECD, 2014). Also, data sharing increases research visibility increases the pace of innovation and facilitates collaborative problem-solving (Adrian *et al.*, 2018; Patel, 2016; William *et al.*, 2019). Additionally, through data sharing it is possible to verify the original analysis, compare results, and limit the unnecessary collection of existing scientific data (Dallmeier-Tiessen *et al.*, 2014; Fecher *et al.*, 2015a; Jao *et al.*, 2015; Mulder *et al.*, 2017). Without proper allowing access to agricultural scientific data important issues such as climate and food security can be impaired (Havens, 2019). This reminds the scientific community to continue strengthening data sharing. It is for this reason that this study was undertaken to enlighten researchers on the need to share their data.

To achieve effective data sharing there is a need to create a better data-sharing environment. Effective scientific agricultural data sharing is achieved when it is a trustful, transparent environment in data sharing, the presence of clear incentives for data sharing, and compliance with agreed legal terms, specifications, and standards for data sharing (European Commission, 2017). Data sharing can take place through some methods. Research data and other related material can be shared with fellow researchers and other research institutions by submitting to the one requesting data upon request (DPRTC, 2019; Kim & Stanton, 2016; Tedersoo *et al.*, 2021). Depositing data into databases and repositories is another method for data sharing (Parr *et al.*, 2017). The most important motivating factor for a scientist to share their data is the perceived benefits attained through data sharing.

As a result, in recent years there have been several initiatives that put more emphasis on scientific data sharing. For example, universities, research institutions, national governments, funding agencies, and journal publishers have been encouraging researchers to open and share their scientific data (Adrian *et al.*, 2018; Gordon *et al.*, 2018; Williams, 2022). The funding agency, research funding institutions, and publisher's policies for PLoS ONE and data sharing frameworks such as the ODIs Data Spectrum, the FAIR principles, Tim Berners-Lee 5 Star Scheme, and DATA-FAIR principles emphasises opening and sharing scientific data (Ali & Dahlhaus, 2022; Boeckhout, 2018; European Commission, 2017; Federer *et al.*, 2018). The availability of timely and reliable agricultural research data provides the base for agricultural research.

In Africa, there have been notable initiatives for data openness and sharing. The African Data Revolution Report of 2018 indicates that open data initiatives were inadequately embraced by most African countries including Tanzania (Van Bell *et al.*, 2018). In Tanzania like in any other developing country, agricultural research data sharing is not effectively established and implemented in agricultural research institutions despite the existence of initiatives to open up different types of data to the public (Katabalwa *et al.*, 2021). Also, the absence of data-sharing repositories in Tanzania makes research institutions practice their data activities without any national coordination (Fosci *et al.*, 2019). However, there are several efforts done in Tanzania to facilitate data sharing. There is a need for the implementation of policies for better timely dissemination of animal

data on research projects based in Kenya and Tanzania (Bett, 2019). The Government of Tanzania through its different agencies has done a lot to open and enable different types of data to be shared with the public. Furthermore, another initiative implemented by the Government to open data and make data discoverable is the establishment of an open data portal, an open Government initiative in 2011 (Agrawal, 2017; Katabalwa *et al.*, 2021; Mboera, 2015; McMurren *et al.*, 2017). These initiatives did not attain effective data-sharing practises in agricultural research institutions however, some remnants still exist.

In the health sector, the health research institutions in Tanzania managed to implement data-sharing policies to guide data management (MUHAS, 2019). Along the same line, Ifakara Health Research Institution launched and developed a data repository to facilitate data management and sharing in the research institution hence benefiting researchers, the Government and students (IHI, 2012). Along the same line, the Government of Tanzania has successfully managed to implement health information systems for improved health data availability and sharing (Nsaghurwe *et al.*, 2021). These health information systems facilitate management and enhance data sharing in the health sector.

The amendment of Acts and the presence of data exchange and sharing guidelines are efforts done to guide the process of data sharing in Tanzania. The Statistical Act CAP.351 R.E. of 2019 was established to facilitate the NBS in the collection, management, and dissemination (URT, 2019b). In line with this, the e-Government Act of Tanzania sections 48, 58, and 59 put the need to create primary data once and share it with the public. While sharing an individual must adhere to the data sharing and exchange principles which ensure effective and efficient data sharing (URT, 2019). Furthermore, the Government has established electronic data exchange and sharing guideline that manages data exchange and sharing to all public institutions in Tanzania (URT, 2019a).

The Tanzania Commission for Science and Technology (COSTECH) guideline for research registration in Tanzania stipulates clearly that researchers are required to share their raw data in soft copy with COSTECH and local collaborating institutions based on project agreement (COSTECH, 2022). The presence of this guideline to an institution coordinating research activities in Tanzania reveals the existence of data-sharing practices between researchers, COSTECH and their research institutions. Also, the research guideline for SUA put the data sharing procedure, and methods throughout the project research lifecycle (DPRTC, 2019). Despite all the efforts done in the implementation of data-sharing initiatives in Tanzania, there is a lack of a clear picture that reveals the current data-sharing practices among researchers, and what motivates researchers to share their data. The current study provides additional insight and understanding of how individual and institutional factors drive researchers' scientific data-sharing practices. The study will add value to the scientific community as it sheds light on the need to establish data-sharing and exchange policies in agricultural research institutions which will provide a formal mechanism for data management and sharing.

1.1 Statement of the Problem

Agricultural researchers are users and also producers of data (Hua and Wang, 2016). Agriculture researchers dedicate most of their time to generating data, and new innovative technologies that facilitate improving crop production (Kahinga, 2014; Ng'eno & Mutula, 2018). Such data generated by researchers are needed to be shared with the scientific community. Prior research from different research disciplines including that of agriculture

has already exposed the methods used in data sharing, what motivates and hinders researchers to share the produced data (Federer *et al.*, 2015; Lodwick, 2019; Moore *et al.*, 2021; Mulder *et al.*, 2017; Tenopir *et al.*, 2018; Williams, 2022; Zuiderwijk & Spiers, 2019), despite all previous literatures and the existence of data exchange and sharing guideline in Tanzania (URT, 2019a), TARI ACT of 2016 (TARI, 2016), COSTECH research registration guideline (COSTECH, 2022) and SUA research guideline of 2019 (DPRTC, 2019). What has been noted from the prior studies and other documented evidences from Tanzania, is that there was no comprehensive look at agricultural research data sharing practices among researchers, it is not clear about the stakeholders with whom agricultural researchers share their data, and the factors that motivate or hinder agricultural researchers in data sharing, particularly in agricultural research institutions in Tanzania. The internal data sharing of research data and some innovative technologies among agricultural researchers may not be similar as only a limited number of research institutions have or may not have an elaborate system for data sharing. With this then what has been observed is that there is a dearth of information on the agricultural research data sharing practises among agricultural researchers, and what motivates researchers in their data sharing in pursuit of improving agricultural research data accessibility.

1.2 General Objective

The main objective of this study was to investigate research data-sharing practices, and what motivates agricultural researchers to share their data in Tanzania.

1.2.1 Specific objectives

- i. Examine the data-sharing practises among agricultural researchers
- ii. Determine factors that influence agricultural researchers in data-sharing practises
- iii. Find out the challenges facing agricultural researchers in their data sharing

1.2.2 Research questions

- i. What are the data-sharing practices exhibited by agricultural researchers in Tanzania?
- ii. Which factors influence agricultural researchers in their data-sharing practises?
- iii. What hinders agricultural researchers in their data sharing practises among researchers?

1.2.3 Significance of the Study

The findings from this study will provide insights for researchers and top-level managers in agricultural research institutions. For researcher's findings will provide an assessment of their contribution to agricultural research data sharing. The knowledge of data sharing will help researchers to improve the data sharing practises in research institutions. The research institution's management will know the limitations researchers are facing in trying to share agricultural research data and may try to solve them. The finding will enlighten Government and research institutions on the need to establish data sharing policy and develop technological infrastructure that could promote and strengthen centralised collaborative data sharing.

2.0 Literature Review

The literature review was conducted based on the study research questions; agricultural researchers' data sharing practices, factors influencing researchers in data sharing, and challenges facing researchers in their data sharing.

2.1 Data Sharing Practices by Agricultural Researchers

Data can be defined depending on the field of study. Data in agriculture include facts obtained through the research process including agronomic samples, satellite photographs, experimental samples, research materials, procedures, technologies, and products (Federer *et al.*, 2015; Kirub, 2016a). Agriculture sciences research data include data obtained from experiments, laboratory or surveys, collections of published peer-reviewed datasets, published research results, satellite photographs, and digital images (Havens, 2019; Kerby, 2015; Kirub, 2016a). Such important data are required to be shared with the scientific community.

Data sharing is an activity carried out by scientists to release their data to other people to use for scientific research (Tenopir *et al.*, 2015). In agricultural sciences, data sharing takes place involving the scientific community to share their common research interests. In a study conducted by Kerby, (2015) and Lodwick, (2019) findings indicate that 20% of data produced were publicly made open through depositing to repositories such as Gene Bank or posting on the website as supplementary files. Furthermore, Williams (2022) reveals that agricultural scientists at the University of Illinois in the USA tended to share their data using three methods; 88% shared their data as supplementary materials on journal articles, 40% shared their data via repositories and 17% shared their data through sending upon request. In line with this, the study that examined the data sharing trends in the College of veterinary medicine faculty members at the University of Oxford in the UK Lodwick, (2019) revealed that 56% of veterinary researchers indicated they shared their data by sending it upon request to fellow researchers, and through publishing articles in journals. Likewise, agricultural researchers in Germany were found to make data available to support research and make data available through a personal request to colleagues in the context of research cooperation (Scenft *et al.*, 2022). In the absence of data-sharing infrastructure, data sharing can be hindered therefore, to strengthen effective scientific data sharing there is a need to establish data-sharing platforms (Huang and Ma, 2015; Li and Zhou, 2008). The presence of a data-sharing platform allows navigation, searching, and downloading of data for use/reuse. Also, this argument has been supported by other prior studies which revealed that to archive effective data sharing there is a need to have data repositories that use common metadata standards as agreed among participating members and adhere to data sharing principles (use of the five-star development scheme and the FAIR Guiding Principles) (Havens, 2019; Late & Kekalainen, 2020; Moore *et al.*, 2021; Tenopir *et al.*, 2015). However, data sharing through depositing to data repositories is the best practice, yet it is not common to most of the agricultural research community, especially in developing countries including Tanzania. But the presence of data repositories is the best practise in data sharing simply because it ensures long-term data availability and accessibility (Moore *et al.*, 2020). In Tanzania for example, there are no national data repositories in place if researchers have their data there is no well-established infrastructure to share data to benefit the scientific community this is the challenge (Mboera, 2015). Therefore, it is not clear the best data-sharing practises exhibited by researchers in Tanzania.

2.2 Factors influencing data sharing behaviours among researchers

Some factors can motivate or hinder researchers to share their data. The perceived benefits of data sharing have a greater role in motivating researchers to share their data. Data sharing helps in answering research questions, testing new hypotheses, serves costs, and efforts to collect data that already exist (Fecher *et al.*, 2015b; Mulder *et al.*, 2017; Pouchard, 2016; Tenopir *et al.*, 2015; Wallis *et al.*, 2013). Similarly, the study that investigated factors that drive or inhibit researchers from openly sharing their data (Zuiderwijk *et al.*, 2020) identified factors that influence a researcher's data sharing including researcher's background, intrinsic motivations, facilitating conditions, trust, expected performance, social influence and affiliation, researchers' experience and skills, and data characteristics.

In addition, the perceived ability to easy-to-access data from a repository motivated researchers in Minnesota to share their data with repositories (Cooper *et al.*, 2017) as well as ensuring to get incentives and provide formal data citations to author's works (Lodwick, 2019; Mulder *et al.*, 2017; Suhr *et al.*, 2020; Tenopir *et al.*, 2015, 2018; Zuiderwijk & Spiers, 2019). However, it was also found that financial gain was not a factor to encourage researchers to share their data because an individual may be willing to share data without expecting any financial gain (Fecher *et al.*, 2017). Equally, an individual researcher's commitment is an important factor influencing researchers to willingly make their data visible (Barakabitze *et al.*, 2015; Bezuidenhout & Chakauya, 2018). Similarly, funding agencies and journal publishers' roles may influence agricultural data-sharing practises. Earlier literature has indicated that conditions from funding agencies, organizations, and journal publisher policies for example PLOS One mandate researchers to open up their data from all publicly funded projects (Kim & Stanton, 2016; Lodwick, 2019; William *et al.*, 2019).

Despite the benefits of data sharing but there are observed reasons hindering researchers in their data sharing. The hypothesized reasons for researchers hesitating to make data available are the; lack of trust, and confidentiality, fear of being scooped, lack of privacy, lack of incentive, lack of money, lack of time, exposure of data to competitors, misuse, and misinterpretation, and lack of rewards (Agarwal *et al.*, 2021; Cooper *et al.*, 2019; Mongeon *et al.*, 2017; Moore *et al.*, 2021; OECD, 2014; Tenopir *et al.*, 2015; Williams *et al.*, 2019; Zuiderwijk *et al.*, 2020). Prior literature on data sharing has shown the following gaps: Data sharing is different across all research disciplines (Mongeon *et al.*, 2017; Tenopir *et al.*, 2015; Zenk-Möltgen *et al.*, 2018). For example, researchers in a human subject such as Medicine/Health sciences were significantly lesser willing to share their data compared to hard sciences including Agriculture and Natural Resources for the reason that it is unethical to share some human-based data (Tenopir *et al.*, 2015).

Most of the prior studies relied on the online survey or desktop research, reviewed several kinds of literature or conducted a meta-analysis reviewing articles to find out data sharing, while other literature used mixed methods to investigate data sharing behaviour on the individual level from different research disciplines (Kerby, 2015; Lodwick, 2019; Mongeon *et al.*, 2017; Park, 2018a, 2018b; Zenk-Möltgen *et al.*, 2018). However, few studies used the cross-sectional survey to investigate data sharing among agricultural researchers (European Commission, 2017; Williams, 2022). What is noted there is a limited study that used the cross-sectional survey to investigate data sharing among agricultural researchers in Tanzania. Other prior work (Zenk-Möltgen *et al.*, 2018; Zuiderwijk *et al.*, 2020)

developed the theoretical research model, to test hypotheses while there are limited previous studies that used TPB to test factors influencing data sharing among agricultural researchers in Tanzania. Based on these findings the current investigation sought to gain a better understanding of what motivates and hinders data sharing by examining both individual and institutional factors.

2.3 Theoretical Framework

This study adopted the Theory of Planned Behaviour (TPB) (Ajzen, 1991). The Theory of Planned Behaviour (TPB) was observed to be a relevant theory to guide this study simply because the theory explains people's behaviours in terms of their attitudes, and resource-facilitating conditions (Zenk-Moltgen, 2018). Theory of Planned Behaviour (TPB) can guide us in understanding the personal and institutional factors that influence or hinder data sharing among researchers (Figure 1) below. In this study the attitude was operationised to include all individual factors applied in this study such as perceived community benefits, perceived behavioural control and others. Institutional factors include all factors that are influenced by the institution in this case a particular research institution. Such factors include institutional arrangements, and procedures. The Theory of Planned Behaviour (TPB) (Ajzen, 1991; Ajzen & Fishbein, 1980) has been used as a theoretical framework to explain how individual and institutional factors can influence data-sharing behaviour. Theory of Planned Behaviour (TPB) has successfully been used to explain an individual's behaviour based on several constructs such as subjective norms, perceived behavioural controls, and attitudes toward behaviour (Yoon & Kim, 2019). Theory of Planned Behaviour (TPB) states that behaviour is a result of a conscious decision to act in a certain way (Zenk-Möltgen *et al.*, 2018). Behaviour intentions are influenced by an individual's attitudes toward a given behaviour, subjective norms regarding that behaviour and perceived behavioural control (Ajzen, 1991). Therefore, behavioural intention refers to a person's aim to perform a given behaviour. While perceived behavioural control is an individual's opinions or perception that one is capable (the belief that one can perform a certain behaviour) or has relevant skills and capabilities to perform a certain behaviour (Zenk-Möltgen *et al.*, 2018).

Likewise, subjective norms can be broadly defined as the social pressure (social influence) that one perceives and is influenced by others to engage or not engage in a certain behaviour (Curty *et al.*, 2017). The weakness of this theory is that it does not explain the influence of political factors in data sharing practises. The role of politics is greater to every aspects of the economy therefore researchers in one way may be influenced by the politics of the country in conducting their research activities. Based on the Theory of Planned Behaviour (TPB) the data sharing conceptual model was developed to explain the relationship between independent variables (individual factors and institutional factors) and dependent variables (data sharing). Independent variables from TPB comprise individual factors (perceived community benefits, perceived career benefits, perceived ability to share, perceived data sharing risk, scholarly altruism and data-sharing norms). The institutional factors include funding agency policies, journal publisher's policies, and facilitating conditions (data sharing infrastructures). The individual and institutional factors were found to influence data sharing (dependent variable).

3.0 Research Method

The study employed a cross-sectional research design where a researcher surveyed observed representative subjects at a specific point in time to examine the data-sharing

practices and the factors that influence researchers in data sharing. The study applied both qualitative and quantitative approaches to data gathering and analysis. The qualitative approach was used to find out the views and insights about data-sharing practices. The quantitative approach was used to investigate the effectiveness of data-sharing practices and factors influencing researchers in data-sharing.

3.1 Population and Sampling

The target population of this research included agricultural researchers from different agricultural research fields belonging to Tanzania Agriculture Research Institution (TARI) centres and Sokoine University of Agriculture (SUA) specifically the College of Agriculture, and College of Veterinary Medicine and Biomedical Sciences. The total population of researchers in the study areas was 527. From this study population, the sample size of 227 was obtained using the formula for sample size calculation (Yamane, 1967).

To obtain the number of respondents to be involved in this study in each research institution a representative sample per institution was calculated by taking the sample size divided by the total population of the study area and multiplied by the population of a particular research institution. The key informant interview was selected purposively including 11 Directors from TARI centres and 1 Director from SUA Post Graduate Studies. Key informants were purposively selected because they were thought to be experienced and knowledgeable and would therefore provide in-depth insights about the study topic.

3.2 Data Collection

The data collection involved the use of multiple data-gathering techniques to examine the data-sharing practice and factors influencing researchers in their data-sharing. The cross-section survey method (questionnaire and interview) and focus group discussions (FGDs) were employed in the study. Quantitative data were collected through questionnaires. The questionnaire in hard copies with both open and closed-ended questions was distributed and administered to 227 agricultural researchers from ten selected TARI centres, and SUA. Out of 227 distributed questionnaires about 204 questionnaires were correctly filled and returned and used in this study. While qualitative data were collected through interviews with key informants and focus group discussions. A total of two focus group discussions with agricultural researchers were conducted at Mikocheni and Ilonga TARI Centres. Each focus group discussion had six participants who were selected from among the survey respondents. These research institutions were selected randomly from among several agricultural research institutions. In addition, secondary data were collected from different sources, including the TARI communication strategy of 2019, the TARI Act of 2016, SUA Research Guideline and Regulation of 2019, and the national electronic data exchange and sharing guideline.

3.3 Statistical Data Analysis

Quantitative data were analysed using Statistical Package of Social Science (SPSS) software version 22. The quantitative analysis allowed a researcher to obtain the frequency of the responses. The descriptive analysis allowed a researcher to describe data sharing practices, and influencing factors for data sharing. While qualitative information collected through interviews, and focus groups were recorded, coded, and categorized into similar themes explaining data-sharing practices among agricultural researchers, and were used in the discussion of findings. The inferential statistics were obtained by using the Binary Logistic Regression technique. This was performed to

determine the presence or absence of statistically significant differences between the independent variable with a dependent variable.

3.4 Measurement of Construct

In testing the existence of relationship or nonexistence of a relationship between the dependent variable (data sharing) and independent variables (set of predictors) (factors influencing data sharing) Logistic regression techniques were used.

4.0 Findings and Discussion

4.1 Demographic Characteristics of Respondents

The survey participants' demographic characteristics were recorded for age, gender, level of education, and work experience. The findings in Figure 1 below indicate the majority are male (67%) and female (33%). These findings imply that majority of agricultural researchers who participated in this study were male. Most of the researchers in this study had an age range from 31 to 50 years of age (69%). This indicates that most researchers were of active age for conducting research in the field. The academic levels for the survey indicated that the majority (82%) of researchers have attained a Master's and PhD level. This implies that most of the researchers in the study area had enough education for conducting research activities. The majority (95.3%) had a work experience of more than (6) years and a few (11.8%) had work experience between 1 to 5 years. This implied that the majority of researchers were experienced in conducting research from different agricultural fields.

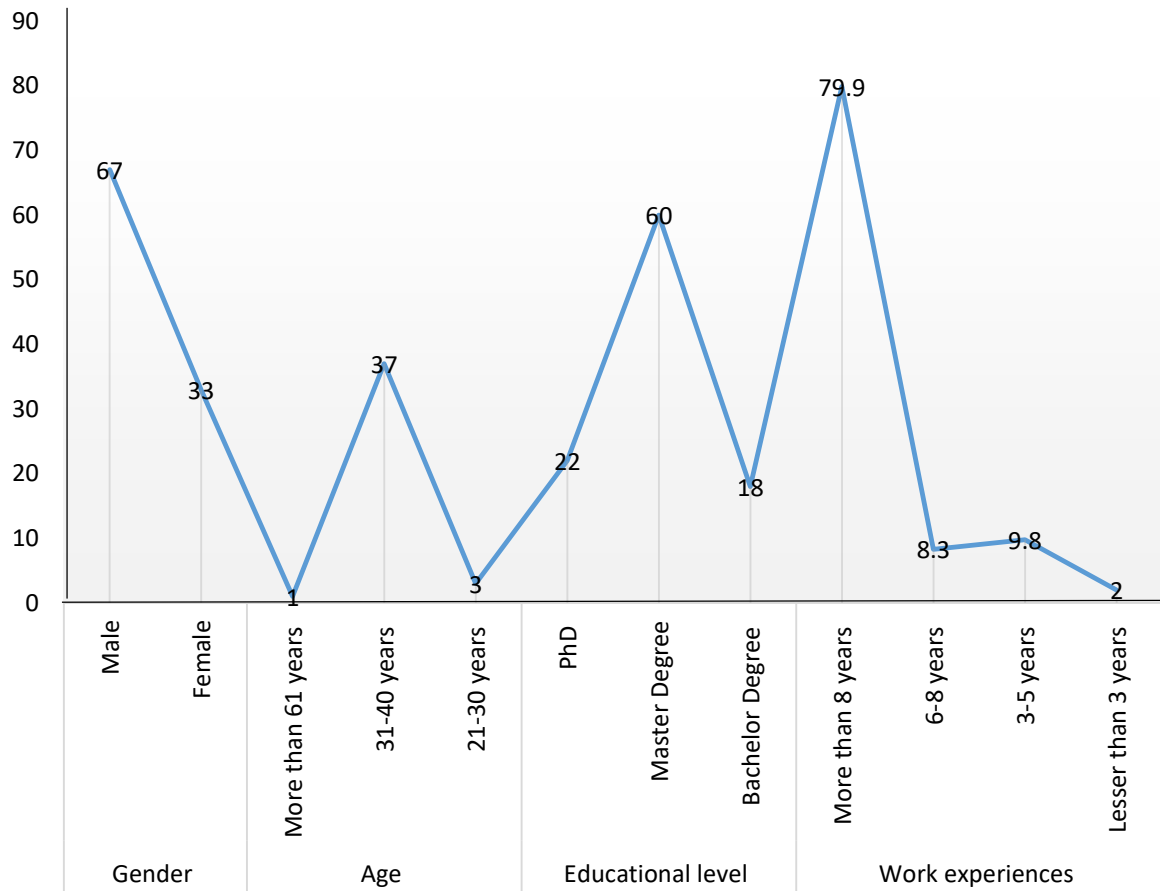


Figure 1: Characteristics of the respondents (n=204)

Source: Field Data, 2021

4.2 Data Sharing among Agricultural Researchers

Findings in Figure 2 below show the responses on data sharing. The question intended to know if respondents share their data. Findings revealed that 74% of respondents indicated they share their data and 25.5% indicated they do not share their data. This finding indicates that the majority of respondents share their data with fellow researchers.

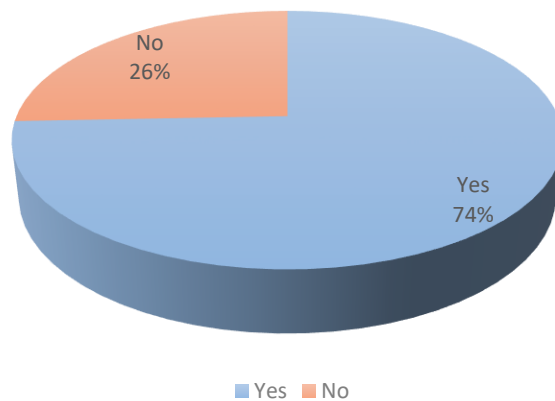


Figure 2: Showing data sharing among agricultural researchers

Source: Field Data, 2021

4.3 Data sharing stakeholders exhibited by agricultural researchers in the study area

Findings in Table 1 indicate that majority of respondents (more than 90%) indicated they shared data with their sponsor of the research project. This means that agricultural researchers preferred to share their data with the project sponsor during and at the end of the research project. This finding is similar to that reported by Shen (2015), who revealed that academic faculty researchers shared their data with funding partners. Also, findings show that more than 80% of respondents indicated that they share their data with members of the research group, shared with the department members, and with their research institution. These findings imply that agricultural researchers work in teams and therefore share their data with members with whom they work closely in their research activities. Therefore, in this way data sharing enhances a collaborative way of working in research. These findings are related to findings of prior studies. Findings from agricultural researchers at the University of Nebraska indicated that they had a tradition of sharing their data with trusted colleagues, a faculty member in another college, other academic institutions, international organizations, and research collaborators (Delserrone & Dinkelman, 2016; Koopman & de Jager, 2016). Furthermore, these findings are related to the previous study that revealed researchers were ready to collaborate and share their data with others. 73% of researchers shared with fellow researchers in the same team, 42% revealed to share their data within the same university, and 55% collaborated with researchers in other institutions (Ünal *et al.*, 2019). In line with this, data sharing enhances interdisciplinary and collaborative research (Aydinoglu *et al.*, 2017). Zhao and Wang (2015) revealed that agricultural researchers were willing to use other researchers' data and are unwilling to provide their data. Through key informants interview one respondent at Ukiriguru TARI centre revealed the following with whom data are shared;

"I am willing to share my data with farmers, project funders, and fellow researchers working on the same project. This is because I have a close connection in research activities with these groups". (Data collected, 2021).

Table 1: Data sharing stakeholders exhibited by agricultural researchers (n=204)

With whom researchers share their data	Yes	
	Frequency	Percentage
Sponsor of research project	196	96.1
Members of my research group	182	89.2
With my department	168	82.4
With my research institution	168	82.4
Journal publishers	154	75.5
With researchers whom we work on similar topic	152	74.5
With researchers whom I know personally	149	73.0
Everyone by publishing online	142	69.6
Institutional library	130	64.0
General public through media	127	62.3

Source: Field Data, 2021

4.4 Factors influencing data sharing practice among researchers

Table 2 shows Inferential Statistics Analysis from Binary Logistic Regression model results. Based on the construct borrowed from TPB in Table 1 above findings indicate that both individual and institutional factors motivate researchers in data-sharing practices. The results reveal that both individual and institutional factors influence researchers in data sharing. The individual factors that were found to have a statistically significant relationship with data sharing were perceived community benefit ($\beta=1.385, p=0.012$), perceived ability to share ($\beta=-1.338, p=0.042$), data sharing norms ($\beta=-500, p=0.050$) and perceived data sharing risk ($\beta=1.625, p=0.015$). While other individual factors including perceived career benefit ($\beta=.274, p=.0607$) and scholarly altruism ($\beta=-0.186, p=0.675$) were found to have no statistically significant relationship with data sharing. At the institutional level, only funding agency policy ($\beta=0.969, p=0.023$) was found to have no significant relationship with data sharing. Other institutional factors such as journal publishers' policy ($\beta=0.089, p=0.836$) and technical infrastructure ($\beta=0.495, p=0.465$) were found to have no statistically significant influence on data sharing. In testing the fitness of the model the Hosmer and Lemeshow Test was found to have a p-value greater than 0.05 ($p=0.131$), indicating that the model was fit for the test. The test Cox & Snell R² was found to have a value of 0.145 and Nagelkerke value of 0.196 meaning that independent variables in the model contributed between 14.5% to 19.6% of the variation on the dependent variable. Thus, the remaining percent was contributed by other factors other than those included in the model. This research demonstrates that agricultural researchers are influenced by both individual and institutional factors in their data sharing. The study, therefore, proposes this model to be used in guiding researchers in data sharing practices Figure 1 above.

Findings indicate that perceived community benefits ($\beta=1.385, p=0.012$) were found to have a statistically significant positive influence on scientist data-sharing behaviour. These findings imply that researchers who perceive that agricultural research data sharing can be beneficial to the community will likely share their data with others. Similar sentiments have been reported by Williams *et al.* (2019), an agricultural researcher from Minnesota revealed share their data by depositing data into repositories they believe that sharing

data through repositories is beneficial to the scientific community whereby researchers can access, browse and download data for future use/reuse in research. Researchers who through data sharing is beneficial to the scientific community as it increases the expectation to collaborate and create a network with fellow researchers on future projects they shared their data (Zuiderwijk *et al.*, 2020).

Scholarly altruism ($\beta=-0.186$, $p=0.675$) was found to have no statistical significance differences between scholarly altruism and data sharing. The Wald value for scholarly altruism was a non-zero value indicating there was a small relationship between scholarly altruism and data sharing but the relationship was not statistically significant. This is contrary to what was found by Kim & Stanton (2016; Williams *et al.* (2019), who found that scholarly altruism had a significant relationship with scientist data sharing. This means that researchers with feelings of scholarly altruism are more likely to share their data than those who did not. Furthermore, perceived behavioural control/perceived efforts ($\beta=1.338$, $p=0.012$) was found to have a statistically significant relationship with scientist data sharing. This means that scientists who have skills, and knowledge and perceive able to share data are likely to share their data. This result supports many prior studies' arguments. Researchers perceived efforts and expected performance was realized to be among the intrinsic and personal factors influencing researchers to openly share data (Zuiderwijk *et al.*, 2020). Also, Kim and Stanton (2016) found that perceived effort had a significant negative effect on scientist data sharing behaviour. This means that scientists who perceive that they require a lot of time in data sharing are less likely to share data with others.

Likewise, data sharing norms ($\beta=0.500$, $p=0.050$) were found to have a statistically significant difference in scientist data sharing behaviour. This means that scientists who perceived that there are data-sharing norms that influence data sharing are likely to share their data. In agricultural research institutions researchers are influenced to share their data by fellow researchers, collaborators, and funding agencies. The TARI Act of 2016 stipulates that all research data and other technology innovated belong to research institutions at the end of the project researchers are required to submit it to the research institutions (URT, 2016). The SUA research guideline also states clearly that all research data and specimens generated by researchers and students at SUA belong to the institution. The guideline further states sharing of data and research materials is upon request by other researchers and universities and research institutions based upon the spirit of collegueship in science and the applicable institutional and national policies and guidelines (DPRTC, 2019). The institutional data exchange and sharing in Tanzania is mediated by electronic data exchange and sharing guidelines (URT, 2019b). All these data-sharing norms in research institutions influence researchers to share their data with the research project funders and institutions. This finding agrees with the prior studies finding argued that data sharing is driven by a peer, the researcher's attitudes, social norms, self-efficient, social, organizational, cultural factors, and disciplinary practices (Cooper *et al.*, 2019; Zenk-Möltgen *et al.*, 2018; Zuiderwijk *et al.*, 2020).

Another individual factor was the perceived risk ($\beta=1.625$, $p=0.015$), which was found to have a statistically significant positive relationship with scientific data-sharing behaviour. This means that scientists who viewed data sharing as a potential risk were lesser likely to share their data and if they thought it was not a risk they could share their data. Similarly, funding agencies' pressure ($\beta=0.969$, $p=0.023$) was found to have a statistically significant

relationship to scientist data sharing. Thus, funding agencies' policies influence researchers in their data-sharing practice. These findings relate to other prior studies that revealed that funding agency policy, and mandates, government agencies, the existence of Government directives was observed to influence researchers in their data sharing (Federer *et al.*, 2015; Higman & Pinfield, 2015; Jao *et al.*, 2015; Park, 2018b; Zenk-Möltgen *et al.*, 2018; Zuiderwijk *et al.*, 2020). However, Kim and Stanton (2016) revealed that funding agency policies were not found to have a significant influence on data-sharing behaviour across diverse STEM disciplines. Also, Lodwick (2019) insists the role of funding agency policies in influencing data sharing; findings showed that journals with no data sharing policy for example *Antiquity* had a lower level of data sharing.

Equally, journal publishers' policies ($\beta=0.089$, $p=0.856$) were found to have no statistically significant difference in data sharing. However, there was some relationship between journal publishers' policies and data sharing because its Wald value was found to be a non-zero value indicating the existence of the relationship. Prior study findings are contrary to these findings by Kim and Stanton (2016) who found that regulative journal pressure positively increased scientist data sharing behaviour.

In the same way, perceived career benefits ($\beta=0.274$, $p=0.607$) were found to have no statistically significant differences in scientist data-sharing behaviour. Although there was no statistically significant relationship the relationship exists to a small extent because the Wald value was larger than the zero-value implying the existence of a minor relationship between perceived career benefits and data-sharing behaviour. These results are related to the previous study (Kim & Stanton, 2016) which found that perceived career benefits were not significantly related to scientist data sharing.

Additionally, facilitating condition (technical infrastructure) ($\beta=0.495$, $p=0.467$) was found to have no statistically significant relationship with scientist data sharing behaviour. The findings reveal there was no statistically significant relationship however, there was the existence of a small relationship between the availability of data infrastructures and data sharing because the Wald value was non-zero indicating the existence of a relationship among the variables. The results support the finding of a prior study which found that the availability of data repositories in a discipline was not found to have a significant relationship with data-sharing behaviour (Kim & Stanton, 2016). Other prior studies revealed that the presence of facilitating conditions (data repositories) can both drive and inhibit researchers from openly sharing their data (Cooper *et al.*, 2019; Thøgersen & Borlund, 2021; Zuiderwijk *et al.*, 2020). Researchers from different research disciplines indicated that they would be willing to share at least some of their data in a central repository if the datasets are easily accessible (Tenopir *et al.*, 2015). In line with these researchers, prior studies found that building a data-sharing infrastructure can promote data sharing (Fecher *et al.*, 2015a; Tenopir *et al.*, 2015). However, these studies reveal that the availability of data repositories is not the only motivator for researchers to share their data, other factors need to be considered.

Table 2: Factors influencing data sharing practice among researchers (n=204)

Factors influencing Data Sharing	B	S.E.	Wald	Df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
							Lower	Upper
Individual factors								
Perceived community benefit	1.385	.554	6.253	1	.012*	3.994	1.349	11.823
Perceived career benefits	.274	.532	.264	1	.607	1.315	.463	3.732
Scholarly altruism	-.186	.444	.176	1	.675	.830	.348	1.982
Perceived ability to share	-1.338	.657	4.149	1	.042*	.262	.072	.951
Data sharing norms (subjective norms)	-.500	.258	3.745	1	.050*	.606	.365	1.006
Perceived data sharing risk	1.625	.667	5.936	1	.015*	5.078	1.374	18.764
Institutional factors								
Funding agency policy	.969	.782	1.535	1	.023`*	2.635	.569	12.201
Journal publishers policy	.089	.488	.033	1	.856	1.093	.420	2.842
Availability of technical infrastructure	.495	.681	.529	1	.467	1.641	.432	6.230
Constant	-.237	.979	.059	1	.808	.789		

Note: * significant at 5% level

Source: Field Data, 2021

4.5 Challenges Hindering Data Sharing among Agricultural Researchers

Despite the fact that researchers indicated willingly sharing their data with fellow researchers, and other colleagues, findings in Table 3 show that 102 (50%) of respondents indicated lack of data sharing policy for guiding data sharing to local and international context was among the factors that hinder researchers in data sharing. The presence of a data-sharing policy could stipulate all important aspects of data sharing including the procedure, methods, and agreements to be considered while sharing data. The presence of policy could influence the development of data repositories to facilitate data-sharing practices. Therefore, the absence of data sharing policy in agricultural research institutions makes it difficult for researchers to recognize a well-elaborated data-sharing scheme. These findings are in line with the study by Katabalwa *et al.* (2021), who revealed that the absence of data sharing policy in Tanzania hampers data sharing. The absence of a data-sharing policy makes it difficult for researchers to be able to share their data with the scientific community. In addition, another prior study revealed that what hinders data sharing includes a lack of data sharing policies, standards, and technical infrastructures such as data repositories (Mboera, 2015; Shen, 2015). Furthermore, findings indicates that insufficient research funds 43.6% was observed to be a factor which hinder researchers in their data sharing. The funds availability allows researchers to conduct research easily thus it could be easy for them to share their data that has been obtained through funded projects.

Table 3: Challenges hindering data sharing among agricultural researchers

Statement	Frequency	Percentage
Lack of data sharing policy	102	50
Insufficient research funds	89	43.6
Lack of reward	74	36.6
Difficult to prepare data to share	73	35.8

Source: Field Data, 2021

5.0 Conclusion

The findings reveal that despite the absence of a well elaborate agricultural research data-sharing policy agricultural researchers share their data with different research groups including project sponsors, members of the research group, departmental members, and research institutions. Findings show that data sharing with research institutions and funding agencies is mandatory. Research institutions and research funding agencies for example COSTECH, and SUA stipulate sharing raw data for all funded projects is mandatory. Researchers are required to submit their raw data and other research materials at the end of the project in both hard copies and soft copies. Findings reveal that both individual and institutional factors motivate agricultural researchers to share their data. For example, for institutional factors funding agency policies and research institutional guidelines requires researchers to share their data and other research materials at the end of the research project. Individual factors also were found to influence researchers in data sharing. It can be concluded that factors that influence data sharing in TARIs and SUA include perceived community benefits, perceived ability to share, and existing data sharing norms such as submitting raw data to institutions, and funding agencies. These would be the factors that respective institutions should promote to improve accessibility to agricultural research data in Tanzania. Although this survey was conducted on a national level in Tanzania, the results of this study can help to address overarching challenges of data sharing at the international level as well. Data sharing is not a national concern, the challenge of data sharing can only be solved in a participatory and cooperative way.

6.0 Recommendations

Based on the study findings it can be recommended that:

- i. To strengthen data-sharing using modern ways there is a need to establish better data-sharing infrastructures such as databases or data banks in agricultural research institutions. Research institutions' authorities should establish a data repository for data management that would enhance data sharing among researchers at local and international level.
- ii. Research institutions should create awareness to researchers about the existence of online data sharing platforms which can provide a room to researchers to be able to share their data with the scientific community.
- iii. Researchers should be informed that it is important to share data to benefit the scientific community rather thinking first about their personal gain.
- iv. Research institutions authorities should establish a data-sharing policy to guide data-sharing practises in agricultural research institutions. The policy should incorporate the existing informal data-sharing practises. This is because the absence of a well elaborate data sharing policy in agricultural research institutions was found to hinder the data sharing practises.
- v. To enhance data sharing among researchers in agricultural research institutions there is a need to address the challenges facing agricultural researchers in their data sharing practises especially provision of enough funds for research activities conducted in agricultural research institutions.

Acknowledgment: Sincere gratitude goes to my study supervisors from the Sokoine University of Agriculture and to my employer Jordan University College for allowing me to pursue further studies at Sokoine University of Agriculture (SUA).

Declaration of research interest: The author (s) declared no potential conflict of interest concerning the research, authorship, and /or publishing of this article.

Funding: The author (s) received no financial support for the research, authorship, or publication of this article.

References

- Adrian, A. M., Emison, B., Musker, R., & Hopkinson, J. (2018). *Open Access & Open Data at PUSH Universities in Partnership with Global Open Data for Agriculture (GODAN)* [<https://doi.org/http://wp.auburn.edu/push/wp-content/uploads/2018/06/GODAN-report.pdf>] site visited on 22/08/2021..
- Agarwal, S., Singh, V., & Gandhi, R. (2021). *Could a Data Sharing Protocol be Agriculture's Missing Link?* Digital Green, a Global Development Organization. [<https://www.thechicagocouncil.org/commentary-and-analysis/blogs/could-data-sharing-protocol-be-agricultures-missing-link>] site visited on 07/08/2022.
- Agrawal, A. (2017). *Data Roadmaps for Sustainable Development: Assessment and Lessons Learned (Issue December)*. [http://www.data4sdgs.org/sites/default/files/services_files/RoadmapsAssessmentReport_Dec2017_final] site visited on 29/10/2022.
- Ajzen, I. (1991). The Theory of Planned Behaviour. *Organizational Behaviour and Human Decision Processes* 50(2): 179–211.
- Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behaviour*. Prentice-Hall.[<https://www.worldcat.org/title/understanding-attitudes-and-predicting-social-behaviour/oclc/5726878>] site visited on 09/07/2022.
- Ali, B., & Dahlhaus, P. (2022). The Role of FAIR Data towards Sustainable Agricultural Performance : A Systematic Literature Review. *Agriculture* 12: 309.
- Aydinoglu, A. U., Dogan, G., & Taskin, Z. (2017). Research data management in Turkey: perceptions and practices. *Library Hi Tech* 35(2): 271–289.
- Barakabitze, A. A., Kitindi, E. J., Sanga, C., & Ayubu, S. (2015). *New technologies for disseminating and communicating agriculture knowledge and information (AKI): Challenges for Agricultural Research Institutes (ARI) in Tanzania. August. Electronic Journal of Information Systems in Developing Countries* 70(2): 1-2.
- Bett, K. (2019). *Agriculture Data Shaping Policy and Changing Lives in Kenya and Tanzania, Retrieved from: [https://www.data4sdgs.org/sites/default/files/services_files/Ag%20Data%20]* site visited on 26/05/2022.
- Bezuidenhout, L., & Chakauya, E. (2018). Hidden concerns of sharing research data by low / middle- income country scientists. *Global Bioethics* 29(1): 39–54.
- Boeckhout, M. (2018). The FAIR guiding principles for data stewardship : fair enough ? *European Journal of Human Genetics* 931–936.
- Cooper, D., Bankston, S., Bracke, M. S., Callahan, B., Chang, H.-F., Delserone, L. M., Diekmann, F., Dinkelman, A. L., Farmer, D., Farrell, S. L., Giovenale, S., Glennie, C., Hanscom, S., Haugen, I., Holdswort, L., Houpert, M., Jenda, C. A., Kerby, E., Kocher, M., Young, S. (2017). Supporting the Changing Research Practises of Agriculture Scholars. *Ithaka S+R*. [https://sr.ithaka.org/wp-content/uploads/2017/06/SR_Report_Agriculture_Scholars_06072017.pdf] site visited on 13/01/2023. [<https://doi.org/10.18665/sr.303663>] site visited on 21/03/2022.
- Cooper, D., & Springer, R. (2019). Data Communities : A New Model for Supporting STEM Data Sharing [Issue Brief]. *Copyright, Fair Use, Scholarly Communication, Etc* (109): 1–28. [<https://digitalcommons.unl.edu/scholcom>] site visited on 20/06/2021.
- Curry, R. G., Crowston, K., Specht, A., Grant, B. W., & Dalton, E. D. (2017). Attitudes and norms affecting scientists' data reuse. *PLoS ONE* 12(12): 1–22.
- Dallmeier-Tiessen, S., Darby, R., Gitmans, K., Lambert, S., Matthews, B., Mele, S., Suhonen, J., & Wilson, M. (2014). Enabling sharing and reuse of scientific data. *New Review of Information Networking* 19(1): 16–43.

- Delserone, L., & Dinkelman, A. (2016). "Investigating the Practises and Needs of Agricultural Researchers at the University of Nebraska-Lincoln." *Faculty Publications, UNL Libraries* 341.[<http://digitalcommons.unl.edu/>] site visited on 06/09/2022.
- DPRTC (2019). *Sokoine University of Agriculture Research Regulations and Guidelines*, (4th .ed.), (Issue June).
- European Commission (2017). *Eip-Agri Workshop Data Sharing: Ensuring Fair Sharing of Digitisation Benefits in Agriculture* (Issue July). [www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/#1389828e7db3] site visited on 23/09/2022.
- Fecher, B., Friesike, S., & Hebing, M. (2015a). *What Drives Academic Data Sharing ?* 1–25. [<https://doi.org/10.1371/journal.pone.0118053>] site visited on 10/03/2022.
- Fecher, B., Friesike, S., & Hebing, M. (2015b). What drives academic data sharing. *PLoS ONE* 10(2): 1–25.
- Federer, L. M., Belter, C. W., Joubert, D. J., Livinski, A., Lu, Y. L., Snyders, L. N., & Thompson, H. (2018). Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLoS ONE* 13(5): 1–12.
- Federer, L. M., Lu, Y., Joubert, D. J., Welsh, J., & Brandys, B. (2015). Biomedical Data Sharing and Reuse : Attitudes and Practices of Clinical and Scientific Research Staff. *PLoS ONE* 10(6): 1–17.
- Fosci, M., Loffreda, L., Chamberlain, A. and Naidoo, N (2019). "Country needs assessment research systems in Tanzania, Report for the SRIA Programme", The UK Department for International Development. [https://assets.publishing.service.gov.uk/media/5ef4ad91e90e075c50609d56/NA_report_Tanzania__Mar_2020_Heart_.pdf] site visited on 21/02/022.
- Gilbert, M., Martie, V. D., & Heila, P. (2019). *Identifying and Implementing Relevant Research Data Management Service for the Library at the University of Dodoma*. *Data Science Journal* 19 (1): 1–9.
- Gordon, A., Al Manir, M. S., Smith, B., Rezaie, A., & Baker, C. J. O. (2018). Design of a framework to support reuse of open data about agriculture. *CEUR Workshop Proceedings* 2275: 1–10.
- Gupta, S., & Muller-Bim, C. (2018). A study of e-Research and its relation with research data lifecycle: A literature perspective,. *Bench making: An International Journal* 25(6): 1656–1680.
- Harper, L., Campbell, J., Cannon, E. K. S., Jung, S., Poelchau, M., Walls, R., Andorf, C., Arnaud, E., Berardini, T. Z., Birkett, C., Cannon, S., Carson, J., Condon, B., Cooper, L., Dunn, N., Elsik, C. G., Farmer, A., Stephen, P., Grant, D., ... Main, D. (2018). *AgBioData consortium recommendations for sustainable genomics and genetics databases for agriculture*. *The Journal Biological Database and Curation* 2018: 1-32.
- Havens, E. (2019). *The Case for Open Data in Agricultural Science*. A capstone project submitted in partial satisfaction of the requirements for the degree of Master of Science in International Agricultural Development, *Davis Winter Quarter*, University of California. [https://iad.ucdavis.edu/sites/g/files/dgvnsk4906/files/inline-files/Erin%20Havens%20_capstone%202019.pdf] site visited on 2309/021.
- Higman, R., & Pinfield, S. (2015). Research data management and openness: The role of data sharing in developing institutional policies and practices. *Program* 49(4): 364–381.

- Ifakara Health Institute (IHI) (2012). *e-Health Open Access repository*, Ifakara Health Institute, Daily News Paper on September 14, 2012.[<https://sites.google.com/a/ihi.or.tz/ihi-main-site/media-room-1/ihi-in-the-news/ihi-in-the-news/openaccessrepositorylaunched>] site visited on 23/09/2022.
- Jao, I., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., Molyneux, S., & Marsh, V. (2015). Research stakeholders' views on benefits and challenges for public health research data sharing in Kenya: The importance of trust and social relations. *PLoS ONE* 10(9): 1–19.
- Kahinga, E. W. (2014). *Knowledge Sharing Practices among Crop Researchers at the Kenya Agricultural Research Institute*, [A thesis submitted in partial fulfilment for the Degree of M Sc. Agricultural Information and Communication Management], Department of Agricultural Economics, University of Nairobi.[[repository.uonbi.ac.ke/bitstream/handle/11295/76273/Kahinga_Knowledge%20sharing%20practices%20among%20crop%20researchers%20at%20the%](http://repository.uonbi.ac.ke/bitstream/handle/11295/76273/Kahinga_Knowledge%20sharing%20practices%20among%20crop%20researchers%20at%20the%20)] site visited on 26/07/2021.
- Katabalwa, A. S., Bates, J., & Abbott, P. (2021). Potential opportunities and risks of sharing agricultural research data in Tanzania Keywords Research environments in Tanzania. *IASSIST Quarterly* 45: 3–4.
- Kerby, E. (2015). Research Data Practices in Veterinary Medicine: A Case Study. *Journal of EScience Librarianship* 4(1): e1073.
- Kim, Y., & Stanton, J. M. (2016). Institutional and Individual Factors Affecting Scientists' Data-Sharing Behaviours: A Multilevel Analysis. *Journal of the Association for Information Science and Technology* 67(4): 776–799.
- Kirub, A. (2016a). *Agricultural Research Data Management: Principles, Policy, and Practice*. Ethiopian Institute of Agriculture Research. [<http://197.156.72.153:8080/xmlui/bitstream/handle/123456789/217/data%20management-book.pdf?sequence=1&isallowed=y>] site visited on 03/02/2022.
- Kirub, A. (2016b). *Agriculture Research Data Management: Principles, Policy, and Practice*. Ethiopian Institute of Agriculture Research. [<http://197.156.72.153:8080/xmlui/bitstream/handle/123456789/217/data%20management-book.pdf?sequence=1&isallowed=y>] site visited on 03/02/2022.
- Koopman, M. M., & de Jager, K. (2016). Archiving South African digital research data: How ready are we? *South African Journal of Science* 112(7–8): 42–49.
- Late, E., & Kekalainen, J. (2020). Use and users of a social science research data archive. *PLoS ONE* 15(8): 1–17.
- Lodwick, L. (2019). Sowing the Seeds of Future Research: Data Sharing, Citation and Reuse in Archaeobotany. *Open Quaternary* 5(7): 1–15.
- Mongeon, P., Robinson-Garcia, N., Jeng, W., & Costas, R. (2017). Incorporating data sharing to the reward system of science: Linking DataCite records to authors in the Web of Science. *Aslib Journal of Information Management* 69(5): 545–556.
- Moore, E. K., Kriesberg, A., Schroeder, S., Geil, K., Haugen, I., Barford, C., Johns, E. M., Arthur, D., Sheffield, M., Ritchie, S. M., Jackson, C., & Parr, C. (2021). Agricultural data management and sharing: Best practices and case study. *Agronomy Journal* October 2020, 1–11.
- Molm, L. (2001). Theories of Social Exchange and Exchange networks. In G. R. and B. S. (Eds) (Ed.), *Hand book of social theory* (pp. 260–272). Sage publishers.
- Mulder, N., Adebamowo, C. A., Adebamowo, S. N., Adebayo, O., Adeleye, O., Alibi, M., Baichoo, S., Benkahla, A., Fadlelmola, F. M., Ghazal, H., Ghedira, K., Matimba, A., Moussa, A., Mungloo-Dilmohamud, Z., Owolabi, M. O., Radouani, F., Rotimi,

- C. N., Stein, D. J., & Souiai, O. (2017). Genomic research data generation, analysis and sharing – challenges in the African setting. *Data Science Journal* 16: 1–15.
- MUHAS. (2020). *Muhimbili University of Health and Allied Science Data Sharing Policy*.
- Mboera, L. E. (2015). *The management of health and biomedical data in Tanzania, Need for a National Scientific Data Policy*. [<https://slideplayer.com/slide/7625850>] site visited on 12/09/2022.
- McMurren, J., Young, A., & Verhulst, S. (2017). *Open Data for Developing Economies Case Studies: Tanzania's Open Education Dashboards*, (Issue July). [<https://odimpart.org/files/case-tanzanias-open-education-dashboards.pdf>] site visited on 13/01/2023.
- Ng'eno, E., & Mutula, S. (2018). Research Data Management (RDM) in agricultural research institutes : a literature review. *Inkanyiso, Jnl Hum & Soc Sci* 10: 28–50.
- Nsaghurwe, A., Dwivedi, V., Ndesanjo, W., Bamsi, H., Basiga, M., Nyella, E., Massawe, J.V., Metzger, J and Taylor, P (2021). One Country's Journey to Interoperability: Tanzania Experience Developing and Implementing a National Health Information Exchange. *BMC Medical Informatics and Decision Making* 21:139.
- OECD. (2014). *Data-driven Innovation for Growth and Well-being: Interim Synthesis Report* (Issue October 2014(1-86)). [<https://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf>] site visited on 10/09/2022.
- Park, H. (2018a). *The Impact of Research Data Sharing and Reuse on Data Citation in STEM Fields*. (Issue December). [Theses and Dissertations], University of Wisconsin-Milwaukee. [<https://dc.uwm.edu/cgi/viewcontent.cgi?article=3010&context=etd>] site visited on 22/09/2022.
- Park, H. (2018b). *The Impact of Research Data Sharing and Reuse on Data Citation in STEM Fields*. [Theses and Dissertations], University of Wisconsin-Milwaukee. [<https://dc.uwm.edu/cgi/viewcontent.cgi?article=3010&context=etd>] sited on 07/02/2021.
- Parliament of the United of Tanzania (2019). The e-Government Act No.10 of 2019, Government Notice No.75 published on 07/02/2020. [<https://www.ega.go.tz/uploads/publications/sw-1581504659-the%20e-government%20general%20regulations,%202020.pdf>] site visited on 21/11/2022.
- Parr, C., Antognoli, E., & Sears, J. (2017). How Agricultural Researchers Share their Data : a Landscape Inventory. *Proceedings of TDWG 1, E20434*: 1–2.
- Patel, D. (2016). Research data management : a conceptual framework. *Library Review* 65(4/5): 226–241.
- Pouchard, L. (2016). Revisiting the Data Lifecycle with Big Data Curation. *International Journal of Digital Curation* 10(2): 176–192.
- Shen, Y. (2015). Research Data Sharing and Reuse Practices of Academic Faculty Researchers : A Study of the Virginia Tech Data Landscape. *International Journal of Digital Curation* 10(2): 157–175.
- Suhr, B., Dungl, J., & Alexander, S. (2020). Search , reuse and sharing of research data in materials science and engineering: A qualitative interview study. *PLoS ONE* 15(9): 1–26.
- Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, A., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K., & Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data* 8(192): 1–11.

- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). *Data Sharing by Scientists : Practices and Perceptions* 6(6): 1–21.
- Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Earth and Space Science Research Data Sharing: Practices and Attitudes of Geophysicists. *Earth and Space Science* 5(12): 891–902.
- Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., & Dorsett, K. (2015). *Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide*. *PLoS ONE* 10(8):e0134826.
- Thoegersen, J. L., & Borlund, P. (2021). Researcher attitudes toward data sharing in public data repositories : a meta-evaluation of studies on researcher data sharing. *Journal of Documentation* 78(7): 1-17.
- URT (2019a). *Electronic data exchange and Sharing*, Dar es Salaam, Presedent Office Public Service Management and Good Governance e-Governemtn Authority. [[https://www.ega.go.tz/uploads/standarddocuments/sw-1640950553-final%20electronic%20data%20sharing%20and%20exchange%20guidelines_edited%20\(1\).pdf](https://www.ega.go.tz/uploads/standarddocuments/sw-1640950553-final%20electronic%20data%20sharing%20and%20exchange%20guidelines_edited%20(1).pdf)] site visited on 28/10/2022.
- URT (2019b). *The Statistics Act, Chapter 351*, 30th November 2019. Accessed at: [nbs.go/nbs/takwimu/Act/statistics-Act-Cap-351](https://www.nbs.go.tz/nbs/takwimu/act/statistics-Act-Cap-351). [[https://www.nbs.go.tz/nbs/takwimu/act/statistics_act_cap_351_\(r.e%202019\).pdf](https://www.nbs.go.tz/nbs/takwimu/act/statistics_act_cap_351_(r.e%202019).pdf)] site visited on 25/10/2022.
- URT (2016). *Tanzania Agricultural Research Institute (TARI) ACT No 10 of 2016*, Dodoma: Tanzania. [https://www.tanzania.go.tz/egov_uploads/documents/sheria_chapa_-Tanzania_Agricultural_Research_inst_sw.pdf] site visited on 13/2/2022 .
- Van Bell, J., Iglesias, C., Mungal, P., Lammerhirt, D., Nuhu, H., Hlabano, M., Nest-Nash, T., & Chaudh, S. (2018). *Africa Data Revolution Report 2018*. The statistics and emerging impact of open data in Africa. [<https://webfoundation.org/docs/2019/03/Africa-data-revolution-report.pdf>] site visited on 08/11/2022.
- Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If We Share Data , Will Anyone Use Them ? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE* 8(7): 1–17.
- William, S. C., Farrell, S. L., Kerby, E. E., & Kocher, M. (2019). Agricultural Researchers' Attitudes Toward Open Access and Data Sharing. *Issues in Science and Technology Librarianship (ISTL)* 91: 1–16.
- Williams, S. C. (2022). Data Practices Ten Years Later: A New Review of Selected Publications by Crop Sciences Faculty. *Journal of Agricultural & Food Information*, 1–37.
- Yamane, T. (1967). *Statistics: An Introductory Analysis*, (2nd .ed.), New York: Harper and Row.
- Yoon, A., & Kim, Y. (2019). The role of data-reuse experience in biological scientists ' data sharing : an empirical analysis. *The Electronic Library* 38(1): 186–208.
- Zenk-Moltgen, W. (2018). 'Factors influencing the data sharing behaviour of researchers in sociology and political sciences. *Journal of Documentation* 74(5): 1053–1073.
- Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V., & Balaban, E. (2018). Factors influencing the data sharing behaviour of researchers in sociology and political science. *Journal of Documentation* 74(5): 1053–1073.

- Zhao, H., & Wang, J. (2015). Development and current situation of agricultural scientific data sharing in China. *IFIP Advances in Information and Communication Technology* 452: 80–86.
- Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyse factors influencing open research data adoption. *PLoS ONE* 1–49.
- Zuiderwijk, A., & Spiers, H. (2019). Sharing and re-using open data: A case study of motivations in astrophysics. *International Journal of Information Management*, 49(6): 228–241.

CHAPTER SIX

6.0 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This chapter gives a summary of the findings, conclusions and recommendations based on the specific objectives of the study.

6.1 Summary of Major Results and Conclusions

This study gathered and analysed data concerning the preservation, sharing, and use/reuse of agricultural research data among agricultural researchers. This study had four objectives which were: (1) to examine the data preservation practises for enhancing agricultural research data reuse among agricultural researchers in Tanzania, and (2) a study on data sources for accessibility and reuse practises among agricultural researchers in Tanzania, (3) to study communication channels and their potential applicability in enhancing agricultural research data sharing among agricultural researchers in Tanzania, (4) to study factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania. Four manuscripts have been prepared from the study each of which is based on a specific objective and each manuscript or article is a chapter in this thesis.

6.1.1 Data preservation practices for enhancing agricultural research data usage among agricultural researchers

The findings on the data preservation practises for enhancing agricultural research data usage among agricultural researchers were discussed in Chapter Two based on Objective One. The findings have revealed that a majority of researchers preferred to preserve their data using different storage devices such as field notebooks, desktop/personal computers, and institutional libraries. Moreover, findings indicate that agricultural researchers preferred to preserve their data for more than six years after the end of the project. The findings also indicate factors that influence researchers in the choice of data preservation methods to be easy to reach, cost-effective storage devices, and support to use the devices, adequate infrastructure for data preservation, and reliable power supply. Based on these findings, it can be concluded that there is yet a great role for research data preservation in enhancing data usage among researchers in Tanzania. It can also be concluded that, in enhancing data sharing and reuse, agricultural research institutions should establish an agricultural research data bank to guarantee the permanent availability of data at all times when needed in research.

6.1.2 Data sources for accessibility and use/reuse practices among agricultural researchers in the study area

The findings of the data sources and accessibility among agriculture researchers were discussed in Chapter Three and were based on Objective Two. The findings indicate that many agricultural researchers use and reuse data that are accessed from different sources. The findings also reveal that researchers use/reuse data as a reference in teaching and learning in research institutions that conduct academic activities. Research data are also used in writing research papers and dissertations as well as in the preparation of technical reports and project proposals. Results from Binary Logistic Regression Analysis techniques findings indicate that independent variables did not have

statistical significance, meaning that all factors used in the model contributed equally to actual data reuse. Based on these findings, it is concluded that agricultural researchers have different sources for data that are used/reused in research activities, and as references in the teaching and learning process.

6.1.3 Communication channels and their potential applicability in enhancing agricultural research data sharing among agricultural researchers.

Chapter Four, which is based on Objective Three, examined communication channels and their potential applicability in enhancing agricultural research data sharing among agricultural researchers. The findings point out that both mediated and non-mediated channels existed and were used as data-sharing avenues. The majority of the researchers preferred to use non-mediated channels. Furthermore, findings show that some mediated data-sharing channels were least used; these include social media platforms, institutional repositories, and funding agency databases. Through cross tabulation, the findings indicate that the extent of data-sharing channel usage at SUA was greater than that at other research institutions investigated. Some channel usage also indicate significant differences, including meetings and conferences $p=0.005$, website $p=0.001$, workshop $p=0.049$, telephone as interpersonal mediated $p=0.045$, and technical report $p=0.045$, meaning that these channels were used across all research institutions. The findings further indicate that more than fifty percent of the respondents agreed that timely delivery, the cost for the channel, and convenience of a channel were among the factors influencing researchers in their channel selection. On the basis of these findings it can be concluded that agricultural researchers use both mediated and non-mediated channels in data sharing at different levels across all research institutions. It can further be concluded that research institutions should invest more in mediated data-sharing channels, for instance databases/data repositories in order to enhance centralised and remotely controlled data-sharing practices among researchers.

6.1.4 Factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania.

Our findings in the factors influencing agricultural researchers' data sharing practises for enhanced accessibility of agriculture research data in Tanzania were discussed in Chapter Five on the basis of Objective Four. Findings reveal that the majority of researchers shared their data with project funders, research institutions, and with department members. Our findings also indicate that few researchers shared their data with the public through media, institutional libraries, and publishing through online journals. The findings indicate that some factors that influence researchers in data sharing appear to have unique influence and were statistically significant at 5%; these factors include, perceived community benefit $p=.012$, perceived ability to share $p=.042$, data sharing norms $p=.050$, perceived risk $p=.015$ and funding agency policies $p=.023$. The findings further indicate that agricultural researchers are facing several challenges in their data sharing including: a lack of data sharing policy that could put clear the procedure, standards, and format, for data sharing among researchers. Another challenge is the absence of data management plans and infrastructure: for example, data repositories that could enhance data sharing and use/reuse practises in research institutions.

6.2 Conclusion

Based on above findings it can be concluded that agricultural researchers preserve their data using different storage devices, researchers also reuse the data accessed from primary and secondary sources. Agriculture researchers share their data through the use of mediated and non-mediated data sharing channels. It can be concluded that agricultural researchers share their data with fellow collaborating members. The main factors motivating researchers in data sharing include perceived community benefits, perceived ability to share and data sharing norms and funding agency policy. It can be concluded that agriculture researchers are facing some challenges in data sharing include a lack of data sharing policy and a lack of data sharing infrastructure.

6.3 Recommendations

Based on the objectives and conclusions of this study, the following are recommendations;

- i. It is recommended that agricultural research institutions should formulate and implement data preservation, sharing, and use/reuse policies, regulations, and establish standards or tools that may guide and facilitate data use/reuse among researchers and other stakeholders to the scientific community. The established policies could guide how agriculture research institutions should preserve, share and reuse the data.
- ii. It is recommended that to facilitate future reuse of data generated by researchers from different research projects, it is recommended research institutions should establish data repositories or databases for preserving all agricultural research data. For example investment in mediated data-sharing channels such as data banks can facilitate a centralised and remote data-sharing and reuse practises among researchers.
- iii. It is recommended that researchers should be willing to allow their data to be reused by others simply because they too admit using or reusing the data that are accessed from other sources.
- iv. It is recommended that research institutions should recognise the role of libraries in facilitating data management and sharing. A library is a place where some of the mediated and non-mediated data-sharing channels are hosted. The data repositories and other print research outputs and other resources embedded with research data can be stored in a library. Therefore, strengthening the existing library infrastructures and building new libraries in agricultural research institutions can have a greater impact on data management sharing and use/reuse practices. The improvement of libraries should go hand in hand with recruiting librarians and ICT technicians who would play a greater role in data management and facilitate data sharing among researchers.
- v. It is recommended that research institutions should strengthen the non-mediated data-sharing channels that are currently used by researchers so that they continue to facilitate data sharing among researchers.
- vi. It is recommended that research institutions should create awareness for researchers to be ready to share their data that are obtained from private project sponsorship. Researchers should be informed that while sharing their data they should think first about the benefit of the community than their gain.

- vii. It is recommended that agricultural researchers should put more effort into exploiting their existing data and using original data during their research activities so that they achieve the best in all aspects of research in different agricultural fields.
- viii. It is recommended that research institutions should work closely with other international research institutions or partnerships so that researchers are allowed to access and share their data with fellow researchers through international web-based channels.

6.4 Contribution of the Study to the Body of Knowledge

The findings of this study will add knowledge to the scientific community especially the data sharing practises exhibited by the developing country like Tanzania. The findings have revealed the current status of data sharing, channels for data sharing, and use/reuse practises among agricultural researchers. This study has practical implications for promoting more reliable, and beneficial data sharing in the agricultural scientific community. The present study has significant theoretical contributions: The present study has significant theoretical contributions: the Theory of Reasoned Action (TRA) has been used to explain factors influencing agricultural researchers in their data reuse practises. Researcher added facilitating condition as among the variable which influence data reuse. The this theory has been extended by adding facilitating condition as an important variable to ensure data reuse is achieved. This study also used the Theory of Planned Behaviours (TPB) to explain and guide researchers on the individual and institutional factors motivating researchers in data sharing behaviour. The Theory of Planned Behaviour has been extended by adding one Perceived community benefit as a factor to influence researchers in data sharing. The Data Life Cycle Model has been used to explain the significance of data preservation in facilitating data sharing and reuse in agricultural research institutions. This model was previously used to explain digital data sharing but the model has been replicated in this study so as to explain the role of data preservation in data life cycle. This study also adopted the concentric Layered Model to explain the most important factors to consider when choosing channels for data sharing. This model was previously used by employees who worked in team in new way of working the model facilitated employees in their choice of channel for information and knowledge sharing. But this study modified this model to suit the variable for this study and it has been used to guide agriculture researchers in their choices of channels for data sharing channels.

6.5 Suggested Areas for Further Research

The current study investigated research data preservation, sharing, and use/reuse in Tanzania's agricultural research institutions with the view of data preservation in enhancing data sharing and use/reuse among agricultural researchers. The current study investigated (10) Tanzania Agricultural Research Institute (TARI) centres and Sokoine University of Agriculture (SUA). However, many other TARI centres deal with animal research, another agricultural institute that focuses on research in forests, and those dealing with other fields such as health research institutes, and fish research institutes. The current study recommends research be extended to cover these research institutes in areas of data management, sharing and use/reuse practises. Furthermore, the current

study should be extended to data sharing and technology transfer from researchers to farmers and farmers to agricultural researchers.

APPENDICES

**Appendix 1: Questionnaire for Agricultural Researchers**

SOKOINE UNIVERSITY OF AGRICULTURE
DEPARTMENT OF INFORMATICS & INFORMATION TECHNOLOGY (DIIT)
COLLEGE OF NATURAL AND APPLIED SCIENCES (CONAS), SOKOINE UNIVERSITY
OF AGRICULTURE (SUA), P.O.BOX 3014, MOROGORO

Dear respondent,

My name is Nolasko Victory Mwinami, a PhD student at the Sokoine University of Agriculture, in the Department of Informatics & Information Technology (DIIT). I am taking this opportunity to request you to participate in this study by filling in this questionnaire. My study is entitled **“Agricultural research data sharing practises among agricultural researchers in Tanzania.”** The information gathered will be treated as strictly confidential and used solely for the research purpose.

SECTION A: DEMOGRAPHIC INFORMATION OF THE RESPONDENT

Kindly select the choice that applies to you by ticking the appropriate option

1. Sex: (a) Male (b) Female
2. Education level: (a) Bachelor Degree (b) Master Degree (c) PhD
3. Age: (a) 21-30 years (b) 31-40 years (c) 41-50 years
(d) 51-60 years (e) Above 61 years
4. Working experience: (a) Lesser than 3years (b) 3-5 years
(c) 6-8years (d) More than 8 years
5. Academic field of specialization.....
6. Institution: a) SUA b) Ilonga c) Naliendele d) Selian
e) Mikocheni f) Uyole g) Tengeru h) Katrin
i) Ukiriguru
7. Zone of your research institution
.....

SECTION B: SOURCES OF AGRICULTURE RESEARCH DATA

Please note that research data can be raw data or original basic data directly produced from the following; lab experiments, surveys, interviews, field trials, human experiences and observations. Research data also include data which are already published in journal articles, reports or in other scientific publications.

8. Do you use agricultural research data? a) Yes b) No

If your answer in question 8 above is "Yes" go to question 10

9. If the answer is No in question 8 explain give reasons why you do not use research data.....

10. Which of the following best describe the sources of agricultural research data you use/reuse?

(Tick all sources that apply)

	Source of agricultural research data	Response
a	Field survey	
b	Field trials	

c	Institutional library	
d	Other research institutions and Universities	
e	Experiment	
f	Fellow researcher	
g	Local database (institutional or departmental database)	
h	Other specify.....	

SECTION C: READINESS TO USE OTHER RESEARCHERS' AGRICULTURAL RESEARCH DATA

11. Have you ever accessed research data from fellow researcher?

a) Yes (b) No

If "Yes" in question 11 above go to question 12, If "No" explain why you did not access data from fellow researchers.....

12. Did you use the accessed data? a) Yes (b) No

If your answer in 12 above is "Yes" go to question 13, If the answer is "No" give reasons.....

13. Which type of agricultural research data do you access and use for your research activities? **(Tick one that apply)**

a) Primary research data b) Secondary research data c) Both primary and secondary research data

14. How often do you access and use agricultural research data from other researchers?

(Tick one that apply) a) Always b) Occasionally c) Very rarely

d) Never

15. How do you use/reuse agricultural research data you access from different data sources? **(Tick all that apply)**

	How data is used	Response
a	Use/reuse data in scientific papers writing	
b	Use/reuse data to answer scientific questions	
c	Use/reuse data to aid interpretation of results	
d	Use/reuse data as reference in training and teaching	
e	Use/reuse in testing theories	
f	Use/reuse in reproducing new study	
g	Use/reuse data in comparing results	
h	Use/reuse data in shaping research questions	
i	Use/reuse data in papers review	
j	Other specify.....	

16. What always motivates you to use/reuse other researcher’s data in your research projects? **(Tick all factors that apply)**

Constructs	Statements	Response
Attitude	Data use/reuse improves my results due to similarity in study purpose	
	Data use/reuse help me answer my study questions	
	Data use/reuse is efficiency due to use of similar tools in new study	
	Data use/reuse save me my money	
Subjective norms (Social influence)	Institutional research arrangement	
	Funders or project collaborators policies	
Facilitating conditions	Availability of quality data from trusted sources	
	Availability of reliable data from trusted data storage infrastructure	
	Availability of accessible data from trusted data storage infrastructure	

SECTION D: AGRICULTURAL RESEARCH DATA STORAGE/PRESERVATION PRACTICES

17. After the research project is over do you store your agricultural research data?

Yes b) No

If the answer in 17 above is "Yes" go to question 18, If the answer in question 17 above is "No" explain why you do not store

18. For how long do you store your agricultural research data? **(Tick one that apply)**

a) Less than a year b) 1–3 years c) 4–6 years d) More than 6 year

19. How do you to store/preserve your research data at your institutions? **(Tick all that apply)**

	Storage Method/Devices	Response
a	Storage devices (Flash sticks, External hard disk, CDs and DVDs)	
b	Use of email	
c	Personal computers desk/laptops	
d	Use lab books	
e	Field notes books	
f	Use of clouds storage	
g	Institutional database/repository	
h	Institutional library	
i	Other specify.....	

20. Which factors motivate you in the choice of data storage/preservation method/devices?

Use the following scale for each statement. **Strongly Agree (SA), Agree (A), Undecided (U), Disagree (D) Strongly Disagree (SD)** (Tick one response for each factor)

Factors Influencing Channel Selection	SA	A	U	D	SD
Device is easy to reach					
Device/method is easy to use					
Device/method is cost effective					
Adequate support to use of method					
Adequate data preservation infrastructure					
Reliable power supply					

SECTION E: AGRICULTURAL RESEARCH DATA SHARING METHODS/CHANNELS

21. Have you ever shared your agricultural research data with others?

a) Yes b) No

If "Yes" in question 21 above go to question 22, If "No" go to question 25 below.

22. Which existing appropriate data sharing channels do you use in sharing your data with other researchers?

Note: Mediated communication channels is a communication that take place via computer networks, platforms or other electronic devices. While non-mediated communication takes place through interpersonal communication **(Tick all that apply)**.

Type	Channels	Responses
Mediated Data Sharing Channels	Email	
	Publishing in online journal platforms	
	Organisation website	
	Telephone (interpersonal mediated conversation)	
	Social media platforms	
	Funding agency database/repositories	
	Institutional repository	
Non-Mediated Data Sharing Channels	Face-to-face discussions	
	Print media publication (technical reports)	
	Face-to-face meeting	
	Face-to-face seminars	

23. What are the factors that motivate you in the selection of the communication channel for sharing agricultural research data? Use the following scale for each statement. **Strongly Agree (SA), Agree (A), Undecided (U), Disagree (D), Strongly Disagree (SD)** **(Tick one response for each statement)**

Statements for choice of channels	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
ICT infrastructure					
Funders policies					

Timely delivery of data					
Need for permanent records					
Convenience to channel					
Data characteristics					

24. With whom do you feel comfortable to share your agricultural research data? **(Tick all that apply)**

	<i>Willing to share my data with</i>	<i>Responses</i>
<i>a</i>	With my department	
<i>b</i>	With my fellow researchers	
<i>c</i>	Members of my research group	
<i>d</i>	Sponsor of research project	
<i>e</i>	With my research institution	
<i>f</i>	Journal publishers	
<i>g</i>	With researchers whom we work on similar topic	
<i>h</i>	With researchers whom I know personally	
<i>i</i>	Everyone by publishing online	
<i>j</i>	Institutional library	
<i>k</i>	General public through media	
<i>l</i>	Other specify.....	

25. State what are the reasons for not being ready to share your agricultural research data? **(Tick all that apply)**

	Reasons for not being willing to share my data	Response
<i>a</i>	Lack of data sharing policy	
<i>b</i>	Lack of effective ICT infrastructures for data sharing	
<i>c</i>	Lack of reward for data sharing	
<i>d</i>	Insufficient of research funds	
<i>e</i>	Do not have enough technical skills	
<i>f</i>	Difficult to prepare data to share	
<i>g</i>	Other specify.....	

SECTION F: MEASURING ITEMS FOR RESEARCH CONSTRUCTS (FACTORS INFLUENCING RESEARCHERS IN DATA SHARING)

26. What motivate you to willing share your agricultural research data? To what extent do you agree on the following statements with respect to willingness to share agricultural research data? **Strongly Agree (SA), Agree (A), Undecided (U), (Disagree (D), Strongly Disagree (SD).** (Tick one response for each statement).

Construct	Statements	SA	A	U	D	SD
	Individual Factors					
Scholarly Altruism	I'm intrinsically motivated to willing share my data					
	I feel comfortable sharing my data					
Perceived Community Benefits	Data sharing improves the quality of research work					
	Data sharing allow validation of research findings					
	Data sharing minimizes duplication of efforts					
	Data sharing improves data usage					
	Data sharing increase visibility of research outputs					
	Data sharing allow re-use or use of existing data					
	Data sharing facilitate to answer research questions					
Perceived Career Benefits	Data citation influence me to share my data					
	Data sharing through articles publication may lead into academic promotion					
Perceived Ability to Share	Assistance provided by experts motivate me to share data					
	My ICT skills and knowledge enables me to share data					
Perceived Risk	Fear that a others can publish data before me					
	Fear on misuse of my agricultural data					
	Fear to expose proprietary technology					
Norms of Data Sharing (social influence)	Fellow researchers influence me to share data					
	Project members influence me to share data					
	Project supervisors influence me in data sharing					
	Research institution influence me to submit data					
	Institutional Factors					
Journal Publishers Policy	Journal publisher's policy influence me to share data.					
	Data sharing is a requirement to some journals					
Funding Agency Policy	Data sharing is a requirement to all publicly funded projects.					

Construct	Statements	SA	A	U	D	SD
	Individual Factors					
Facilitating Conditions	Availability of data sharing infrastructure influence me to share data.					
	Stable internet connectivity motivate me to share data					
	Reliable power supply influence to use data sharing devices hence share data					

G: Recommendations

27. In your own words what can be done to improve data sharing practices among agricultural researchers

.....

Thank you for your time

Appendix 2: Interview Guide

I would appreciate if you could spare some few minutes for an interview with you.

The important information you deliver will be treated confidentially and used for academic purposes.

1. Sex (Please tick the applicable)
 - i. Male
 - ii. Female
2. Education qualification
 - i. Bachelor Degree
 - ii. Master Degree
 - iii. PhD
3. How long have you worked at this research institution
 - i. Lesser than 3 years
 - ii. 3-5 years
 - iii. 6-8years
 - iv. More than 8 years
4. Data sources and use/reuse practices
 - i. Do you use agricultural research data in your research activities? If Yes what type of agricultural research data do you access, generate and use?
 - ii. For which purposes you use/reuse the accessed, and generated agricultural research data?
 - iii. If you have ever used any research data and other technologies in your research, what do you believe to be the benefits achieved?
 - iv. What are the main sources of agricultural research data you normally use/reuse for scientific process?
5. Data storage/preservation practices
 - i. How do you store/preserve your agricultural research data?
 - ii. Where do you preserve your agricultural research data beyond your project lifecycle?
 - iii. What challenges you face in the storage/preservation of research data?
6. Data sharing channels
 - i. How can you describe the existing data sharing channel at your research institution?
 - ii. How do you share your data with project funders and collaborators?
 - iii. Do you use cell phone and other social media as a channels for exchanging data, technologies and other innovation?
7. Data sharing practices
 - i. Have you ever shared your agricultural research data with any one?
 - ii. With whom do you feel willing to share your agricultural research data?
 - iii. What motivate you to be willing to share your research data?
What is your main reason for not being willing to share your agricultural research data?
 - iv. Why it is a requirement to share technical reports and other research materials to be shared at your research institution?

Appendix 3: Focus Group Discussion Guide

1. Do you know the concept of agricultural research data use/reuse?
2. From which sources do you access some important agricultural research data to be used in your research process?
3. What type of agricultural research data do you access, and use/reuse?
4. For what purposes do you use/reuse the agricultural research data accessed from other sources?
5. Do you have data repository or data storage database? If no where do you preserve your agricultural research data beyond your project lifecycle?
6. How do you share your data with project funders and collaborators?
7. Have you ever shared your agricultural research data with any one?
8. How do you describe the existing data sharing channel used by researchers at your research institution?
9. Why it is a requirement to share technical reports and other research materials with your research institution?
10. With whom do you feel willing to share your agricultural research data?
11. What motivate you to be willing to share your research data?
12. What are the challenges that hinders you not being able to share your data?

Appendix 4: Research permits letters**TANZANIA AGRICULTURAL RESEARCH INSTITUTE**

DL: 255 (026) 2961993
 GL: 255 (026) 2961994
 Email: info@tari.go.tz/dg@tari.go.tz
 Website: www.tari.go.tz

In reply please quote.

Ref. No. 269/320/01/04

15th March, 2021

To: TARI CENTER DIRECTOR AND MANAGERS



Makutupora HQs, Arusha Road,
 P.O Box 1571,
DODOMA

RE: INTRODUCING MR. NOLASKO VICTORY MWINAMI

Reference is made to the above caption.

Mr. Nolasko Victory Mwinami is a PhD Student at Sokoine University of Agriculture (SUA) whose research title is "**To investigate Agricultural Research Data Sharing Practices Among Agricultural Researchers in Tanzania**".

He has been granted permission by SUA to undertake his research effectively from March 2021 to March 2022. TARI centres are among of his study areas for successful completion of his PhD.

You are kindly requested to accord him needed support to facilitate gathering of relevant information from your respective centres.

I look forward to your usual cooperation.

Lt. 9

Dr. Geoffrey S. Mkamilo
DIRECTOR GENERAL

DIRECTOR GENERAL
 TARI AGRICULTURAL RESEARCH INSTITUTE
 P. O. Box 1571
 DODOMA

/ copy: Mr. Nolasko Victory Mwinami

CLEARANCE PERMIT FOR CONDUCTING RESEARCH IN TANZANIA



UNITED REPUBLIC OF TANZANIA
MINISTRY OF EDUCATION, SCIENCE AND TECHNOLOGY.
**SOKOINE UNIVERSITY OF AGRICULTURE OFFICE OF THE VICE-
 CHANCELLOR**



P.O Box 3000, CHUO KIKUU, MOROGORO, TANZANIA. Phone:
 +255 (023) 2640006/7/8/9, Direct Line: +255 (023) 2640015,
 E-mail: vc@sua.ac.tz, Website: <https://www.sua.ac.tz>

Our Ref: DPRTC/PLS/D/2017/0001/27

Date: 15th March, 2021

The General Director,
 Tanzania Agricultural Research Institute,
 Makutopora Head Quarters,
 Arusha Road,
 P.O. Box 1571,
DODOMA.

RE: UNIVERSITY STAFF, STUDENTS AND RESEARCHERS CLEARANCE

The Sokoine University of Agriculture was established by University Act No. 7 of 2005 and SUA Charter, 2007 which became operational on 1st January 2007 repealing Act No. 6 of 1984. One of the mission objectives of the university is to generate and apply knowledge through research. For this reason the staff and researchers undertake research activities from time to time.

2. To facilitate the research function, the Vice Chancellor of the Sokoine University of Agriculture (SUA) is empowered to issue research clearance to staff, students, research associate and researchers of SUA on behalf of the Tanzania Commission for Science and Technology.

3. The purpose of this letter is to introduce to you **Mr. Nolasko Victory Mwinami** a bonafide **PhD (Information and Records Studies)** student with Registration number **PLS/D/2017/0001** of SUA. By this letter **Mr. Nolasko Victory Mwinami** has been granted clearance to conduct research in the country. The title of the research in

question is "To Investigate Agricultural Research data sharing Practices among Agricultural Researchers in Tanzania".'

4. The period for which this permission has been granted is from **11th** to **30th March, 2022**. The research will be conducted in **Dodoma regions**.
5. Should some of these areas/institutions/offices be restricted, you are requested to kindly advise the researcher(s) on alternative areas/institutions/ offices which could be visited. In case you may require further information on the researcher please contact me.
6. We thank you in advance for your cooperation and facilitation of this research activity.

Yours sincerely,
Prof. Maulid 'vV. Mwatawaia
FOR: VICE-CHANCELLOR



- c.c. Director, DPRTC, SUA - To note in file.
C.C. Student - Mr. Nolasko Victory Mwinami