RESEARCH DATA MANAGEMENT AND CURATION GUIDELINES

MAY, 2022

PREAMBLE

Egerton University strives to transform lives through quality education and research in exercising its mandate stipulated in the Universities Act No.42 of 2012. Research engagement is a Strategic Goal two and is a core mandate and function of the University exercised in creating, protecting and utilizing research data for academic excellence and innovation among staff, students and advancement of the community. The University strategic plan of 2018-2023 provides the following strategic objectives for research and data management:

SO6: To increase research outputs and disseminate findings
SO7: To increase registrable innovations
SO8: To engage in public policy analysis.
SO9: To engage in research, consultancy and community outreach
SO10: To undertake monitoring and evaluation of research
SO11:To improve University and community environmental management capacities.

The University encourages collaborative research and development through partnerships with national, regional and international institutions and organizations. The University supports research by providing competitive funds for research and seeking funding from external funders to supplement its contribution. The data generated from research ought to be securely stored, visible to other researchers and information disseminated to the community at large. The University guided by its vision, mission and core values, is committed to ensuring that research outputs and new knowledge generated are widely disseminated and accessible. The visibility and verification of research data can be achieved when open access principle is embraced. There is need for regulation of the publications in terms of protecting the integrity/credibility of the University and recognition of the University's role. This guideline provides the framework for research data management to staff and students to protect the integrity/credibility of research knowledge disseminated. The guidelines shall be used together with the Egerton University Statutes, Strategic Plan, University Intellectual Property rights and publication policies and other relevant laws of Kenya.

The Guiding Principles

Vision

To be a world class University for the advancement of humanity

Mission

To generate knowledge and offer exemplary education and training to society for national and global development

Philosophy

Egerton University innovatively influences human development through generation, acquisition, preservation and dissemination of knowledge and skills in Agriculture and related disciplines, guided by the core values

Core Values

National unity and social fairness, Integrity, transparency and accountability Professionalism, Internationalism, Passion for excellence and devotion to duty, Teamwork, Passion for environmental conservation,
Innovativeness and creativity

TABLE OF CONTENTS

1: DEFINITIONS

In this document, the following terms are defined as follows:

1. **Access Restriction:** means tagging of active and stored data with the aim of limiting their processing at any time.
2. **Conflict of Interest (COI)**. Any situation in which real or perceived interests of an individual may run counter to the interests of the Institution or negatively affect their employment or duties.
3. **Data Controller:** means a natural or legal person, public, authority, agency or other body which alone or jointly with others determines the purpose and means of processing of personal data (Data Protection Act No. 24 2019 (I)(2)).
4. **Data curation:** is the management of active data throughout its lifespan and after research for reuse over time by researchers and other interested persons and for educational use in all academic disciplines.
5. **Data Management Plan (DMP):** means a formal document that describes how data will be handled both during the research and after the research is completed.
6. **Data management:** means the administrative process by which the data is acquired, validated, processed, stored and protected throughout its lifespan for its accessibility, reliability, and timeliness to satisfy the needs of the data users.
7. **Data Processor:** means natural or legal person, public authority, agency or other body which processes personal data on behalf of data Controller (Data Protection Act No. 24 2019 (I)(2)).
8. **Intellectual Property (IP)**. All outputs of creative endeavour in any field at the Institution for which legal rights may be obtained or enforced pursuant to the law.
9. **Intellectual Property Rights (IPRs)**. The proprietary rights that may be granted for an invention, mark, design, plant variety, or other type of IP, should the statutory requirements for protection be met to result in a patent, trademark, registered design or plant breeders' right, respectively.
10. **Metadata:** means structured information that describes, explains,

locates, or otherwise makes it easier to retrieve, use, or manage an information resource". Examples of key metadata elements are: Title of the article/document; Author (Creator) of the document; Description of the content; Source of the document, Date created.

11. **Open Access:** means availability of research outputs that can be accessed by anyone in the world and is free of restrictions. This should adhere to General Exemptions of Data Protection Act No. 24 2019 Part VII.
12. **Open Data:** means data that can be freely accessed for validation, reuse and redistribution.
13. **Patent**: A patent is a type of intellectual property that gives its owner the legal right to exclude others from making, using, or selling an invention for a limited period of time in exchange for publishing an enabling public disclosure of the invention.
14. **Research Data:** are defined as documented records which may be numerical scores, textual records, images or audio all of which may be in electronic form used as sources for scientific research that are acceptable in the scientific community for validation of findings.
15. **Researcher:** means a person who is engaged with Egerton University in an academic position, full-time, part-time, whether on full appointment or joint appointment, including honorary and affiliate appointments and assistantships.

2: SCOPE OF THE DATA MANAGEMENT GUIDELINE

2.1 This guideline is formulated within the scope of Kenya Data Protection Act No. 24 2019, Egerton University Research Policy, (*EU/RE/POL/01A*), Intellectual Property Rights Policy **(***EU/RE/POL/01D***,** Publication Policy (*EU/RE/POL/1E*) and Research and Ethical Clearance ( ).

The nature of data within the scope of this guideline includes:

2.2 Associated metadata that directly support or substantiate published research findings. Such data are required for validation, and should be made available concurrently with the research publication.

2.3 Valuable data assets that are created by research project(s), but not directly underscoring published research findings. Such data should be made available within the time period indicated by normal practice in the discipline.

2.4 A data management plan embedded in the project proposal, to describe

the most significant and potentially valuable data assets to be generated (ANNEX 1). This should not preclude the sharing and retention of unanticipated but valuable data assets.

2.5 Availability of Software and Tools such as code, algorithms, protocols, analytical and visualization tools required to generate and replicate the data. These components are essential resources for deposition of data.

2.6 This document may be cited as Egerton University Research Data Management and Curation Guidelines.

3: OBJECTIVES OF THE GUIDELINES

These guidelines aims at enhancing documentation, protection, integrity and accessibility to research data for meta-analysis. The objectives to achieve this are:

3.1 Use active and archived data for achieving the Sustainable Development Goals as per Kenya's Vision 2030.

3.2 Create a framework for research data sharing.

3.3 Ensure researchers adhere to ethical principles and practices of data sharing.

3.4 Ensure researchers adhere to integrity of practice that is related to data management principles that support data sharing in a discoverable forum and reusable for specific purposes.

3.5 Enhance transparency and accountability of accurate data collection and analysis.

4: LIMIT OF OPENNESS

4.1 Protection of personal data is within the Data Protection Act No. 24 of 2019

4.2 Commercial interests and security concerns are universally recognized as necessary limits on data availability.

4.3There should be restriction of the availability of data from private and publicly funded projects as stated in the Data Protection Act 2019.

4.4 Ongoing research should not be compromised by premature data release.

5: DATA MANAGEMENT PRINCIPLES

Socioeconomic development within Kenyan communities and beyond

depends on appropriate, pragmatic and result researches with accrual benefits for all sectors in the country. Data are important before and after the life span of research projects. The following management principles of managing research data provide a transparent and coherent framework to this policy:

5.1 All research data that originate from external/public funded research projects linked to or funded from exchequer are public outputs produced for public consumption hence should be available with few restrictions.
5.2 Researchers should be entitled to a limited period of privileged use of the data they have collected to enable them to publish the results of their research. The time frame may be determined by the funding entity/ and should not exceed a maximum period of 2 years following the date of publication.
5.3 Long-term value data should be preserved and made accessible for future research use.
5.4 Generated data are owned by Egerton University unless agreed to by the University in terms of a funder contract.
5.5 Metadata should be documented and made openly available to enable other researchers make references to the research data and re-use them. Visibility of data allows accessibility and availability of research information.
5.6 Policies, guidelines, and practices should be followed to avoid damage due to inappropriate handling and release of data during the life span of research project.
5.7 In respect to intellectual contributions of researchers who generate and preserve data, users of such data should acknowledge the appropriate sources and adhere to its terms and conditions including avoidance of plagiarism and duplication of research.

6: ETHICS AND INTEGRITY OF RESEARCH DATA

Ethical consideration in data management spans across the entire process of research from ethical clearance, data collection (accurate representation), storage and security (preventing loss of data), data transfer (accurate recording and transfer of data), sharing of data and use of data (analyses). It is necessary to maintain high integrity of research conducted and data management should conform to Egerton University Research Policy

(*EU/RE/POL/01A*,) and Standard Operating Procedures for Egerton University Research Ethics (*EUREC*).

.

6.1 Principal investigator is responsible for ensuring the integrity of the data as described in University Research Policy. However, by extension and in view of importance of research data, this responsibility extends to all persons involved in planning the study, collecting the data, analyzing or interpreting the research findings, publishing the results of the study, or maintaining the research records.

6.2 The integrity of research depends on all aspects of data management, including but not limited to collection, ownership, use, storage, and sharing of data. Researchers are therefore expected to uphold standards of good research practices in the design, conduct and reporting results.

6.3 Fabrication and falsification of data are major threats to data integrity. An independent researcher should be able to repeat the experiment and analysis to verify validity of the data.

6.4 Verification of results/outputs is important for ethical and assurance of the public on the trust bestowed upon researchers. As a result, open data enables other scientists to validate research results and also avoid duplication of research but build on the research further.

6.5 Validation of data gives an opportunity to national, international, private and public organizations to engage in a transparent manner.

6.6 Data generated in collaboration with external funds or from publicly-funded research are for public consumption and should be readily available free of charge for reuse.

7: PERIOD OF EXCLUSIVE USE OF DATA

7.1 A privileged period of exclusive use is recognized in line with acceptable limits of openness, after 2 years following the date of publication.

7.2 Restricted period should not apply to data that support published findings and which are necessary for validation.

7.3 User requesting access to restricted data must agree to data use agreement with the data controller.

8: PROCEDURES OF IMPLEMENTATION

8.1 Criteria for Selection of Research Data

Sound Research data and findings are valued because of their originality and contribution to the scientific knowledge and community to solve existing problems for advancement of socio-economic and geo-political development. Some research data are generated using unique equipment/ procedures and may not be possible to repeat in other environments or places. Such unique data may be of importance and valuable to other researchers worldwide. Consequently, some of the essential criteria of data to be considered for management are:

8.1.1 Data generated from the project should be consolidated irrespective of the nature.
8.1.2 Data that may be difficult to repeat should be stored in a repository.
8.1.3 Data that support research findings but may not have been used in publication. Such data should be stored for further verification.
8.1.4 Data from longitudinal studies of human or natural events that may not be possible to repeat on the same individuals and require strict observation of ethical issues.
8.1.5 Experimental results such as specimen, samples, germplasm that would be expensive to reproduce.

8.2 Role of Researcher

The roles of researcher in implementation of data management are to:

8.2.1 follow funding requirements specified in proposal and grant agreement, during and after the active lifespan of a research project.
8.2.2 describe data to be generated from the research project as required for in data management plan (ANNEX 1).
8.2.3 describe how data shall be stored and managed according to the requirements in the plan.
8.2.4 Submit proposal that describes how data shall be managed for transmission to the funding entity.
8.2.5 Store, manage and be custodian of the data during the life span of the project as specified by the funding entity.
8.2.6 Submit research data timely to Egerton University, Research and

Extension Division repository or domain-specific data infrastructures according to funding entity requirements.

8.2.7 deposit research notebooks in form of hard or soft copies with Egerton University at the appropriate time.

8.2.8 operate within the Researchers' rights and responsibilities as indicated in the Egerton University Research Policy EU/*RE/POL/01A*.

8.3 Role of Egerton University

8.3.1 Members of Egerton University Research Committee shall manage data as stipulated in Research Policy EU/*RE/POL/01A*, **clause 24**.

8.3.2 Adhere to the requirements and policies of the funding entity.

8.3.3 Create awareness among researchers and academicians on policies and guidelines pertaining to data management practices.

8.3.4 Ensures that compliance with policy requirements by grant holders is adequately supported.

8.3.5 Describe and maintain research data management guidelines at institutional level.

8.3.6 Enable management planning and execution of good data management practice by specific research projects.

8.3.7 Maintain the raw data before summary, analysis and reporting. This is essential for validation and verification.

8.3.8 Ensure that data are visualized and accessed accordingly.

8.3.9 Provide secure long-term storage and maintenance of data as per Egerton University Research Policy *EU/RE/POL/01A*, and Data Protection Act No. 24, 2019.

9: MANAGEMENT OF RESEARCH DATA BANK

In research, data management is part and parcel of research activity, and it is appropriate to secure and use public funds to support management of data. Egerton University, Division of Research & Extension is responsible for long-term storage and maintenance of data (Egerton University Research Policy, *EU/RE/POL/01A*,). Therefore, all proposals should include cost of managing data.

Data are managed/stored at two levels:

9.1 Researchers Level

9.1.1 It is the responsibility of researchers, including postgraduate and undergraduate students undertaking research, to manage effectively the data they create and submit the same to the university in a prescribed manner.
9.1.2 Researchers shall declare their practice as described in the management plan. Egerton University shall monitor and evaluate the execution of the research as stated in the Egerton University Research Policy (*EU/RE/POL/01A*).
9.1.3 Researchers shall store raw and analyzed data in both electronic and hard copy formats in such a way to consider ethical issues, integrity, and ease of access.
9.1.4 Special consideration may be required, especially if a marketable product is produced or in the case of change of institution by the Principal Investigator which may require negotiation and shall conform with Egerton University Intellectual Property Rights Policy.

9.2 Institutional Level

9.2.1 To accomplish efficient data storage management, appropriate software(s) for data input, management and for long term storage shall be installed.
9.2.2 Data storage may not be the same for all areas of research. However, research data should be stored according to the specific discipline. The software installed shall be versatile to accommodate several disciplines.
9.2.3 All researches shall be required to follow and use a data management plan and appropriate software (ANNEX 1).

10: VISIBILITY AND UTILIZATION OF DATA

10.1 Visibility of Research Data

Research project generate raw data that must be stored, analyzed and synthesized into appropriate dissemination products. To increase visibility of Egerton University, the refined data should be stored electronically in an easily accessible format.
10.1.1.All project data shall be assigned a number for ease of identification

and visibility

10.1.2 Research data which may be in raw form should be made accessible to other users upon publication of the manuscript or after at least two years after completion of the project.

10.1.3 Upon submission of a manuscript for publication, Researchers shall channel manuscript through R&E division for assignment of a technical number for documentation, accountability and visibility. The Technical number will be an identifier of Egerton University, Faculty, Department and Year. For example, (EUAGCHS20-1 - where EU - Egerton university, AG - Faculty of Agriculture, CHS is Department of Crops, Horticulture and Soils), 20-year of publication, 1-series of publication. Once the manuscript is published, it will be attached to the relevant Digital Object Identifiers (DOIs).

.

10.2 Sharing of Data and Information

10.2.1 Sharing of data is an acceptable good practice in research provided that the risks such as loss of credit, prejudiced scrutiny of competitors/ detractors, compromising confidentiality of human participants and that the overall value of such exercise will not be lost.

10.2.2 Sharing of data shall be guided by a formal data transfer and sharing agreement.

10.2.3 Shared data shall be classified into identifiable, potentially identifiable and anonymous data.

10.2.4 Sharing of personal data shall be avoided and if necessary, such data must be identified.

ANNEX 1

Egerton University
DATA MANAGEMENT PLAN

	Description
Administrative Data	
Donor/Funder	
Grant Contract Number	
Project Name	
Description of Project	
Name of PI/Researcher	
Egerton University	
Faculty	
Department	
Project Data Contact	
Date ff first Version	
Date of Last Version	
Related Institutional Policies	
Donor Policy	
Ethics and legal Compliance	
Ethical issues	

Copy rights and Ownership	
Period of Access Restriction	
Participants Protection	
Data	
Date of Data collection:	
Date of data Entry	
Method of data Collection	
Format of the data	
Quantitative data	
Qualitative data	
Metadata	
Software Used to analyze Data	
Analyzed data (including links)	
Metadata accompanying the data	
Innovation/Technology	
Publications from the research data	
Storage and Backup	
Egerton University Repository	
Capacity of the Server	
Frequency of backing up data	
Management of data security	

Method of transmitting data to the server	
Long term data storage plan	…
Capacity of Server:	
Cost of server	
Maintenance and management cost	
Data Sharing	
Data sharing platform Egerton University Website	
Duration of Restrictions of Data Sharing	
Duration of minimum data sharing	
Agreement of Data Sharing	
Responsibilities and Resources	
Principal Investigator Egerton University Data Manager	
Qualified staff	
Hardware	
Versatile Software	