RESEARCH DATA CURATION AT KENYA'S AGRICULTURAL RESEARCH INSTITUTE LIBRARIES: OPPORTUNITIES AND CHALLENGES

NG'ENO, Emily Jeruto
ngenoemily28@gmail.com
engeno@mu.ac.ke

ODERO, Damaris JN
oderodjn@gmail.com
oderodjn@mu.ac.ke

AMOTH, Duncan
duncanamoth@yahoo.com
duncanamoth@mu.ac.ke

INTRODUCTION

RESEARCH DATA CURATION (RDC)

The active and ongoing management of data through its lifecycle of interest and usefulness.

Horowitz (2019)

RESEARCH DATA MANAGEMENT (RDM)

Process consisting of data creation, storage, security, preservation, retrieval, sharing, and reuse while accounting for technical capabilities, ethical considerations, legal issues, human resource capability and government frameworks

Ray (2014); Whyte and Tedds (2011)

Significance Of Agricultural Research Data Curation:

- Improving the Quality of Research Reproducibility

- Enhances Usability and Accessibility

- Scrutinises agricultural research findings;

- Encourages data sharing and collaboration;

- Supports informed decision-making;

- Increase the impact and visibility of agricultural research;

- Enables long-term preservation

- Promotes Compliance and Ethical Standards.

- Economic Impact

(Lewis, 2010; and Van den Eynden, et al, 2011)

Problem Statement

- Agriculture is the bastion of Kenya's economy

- The Kenya Agricultural and Livestock Research Act No.17 of 2013 is silent on RDM

- Agriculture research data curation challenges are known yet remain unanswered

- Results - valuable datasets loss, low sharing and exchange, low quality of research outputs, duplication of research, high data gathering costs, poor re-analysis of existing research data.

STUDY PURPOSE

- Examine how Kenya's agricultural research institute libraries curate their research data;

- Propose how the institute libraries can realign themselves to offer re-purposed data curation services.

DATA CURATION CENTRE (DCC) LIFECYCLE MODEL
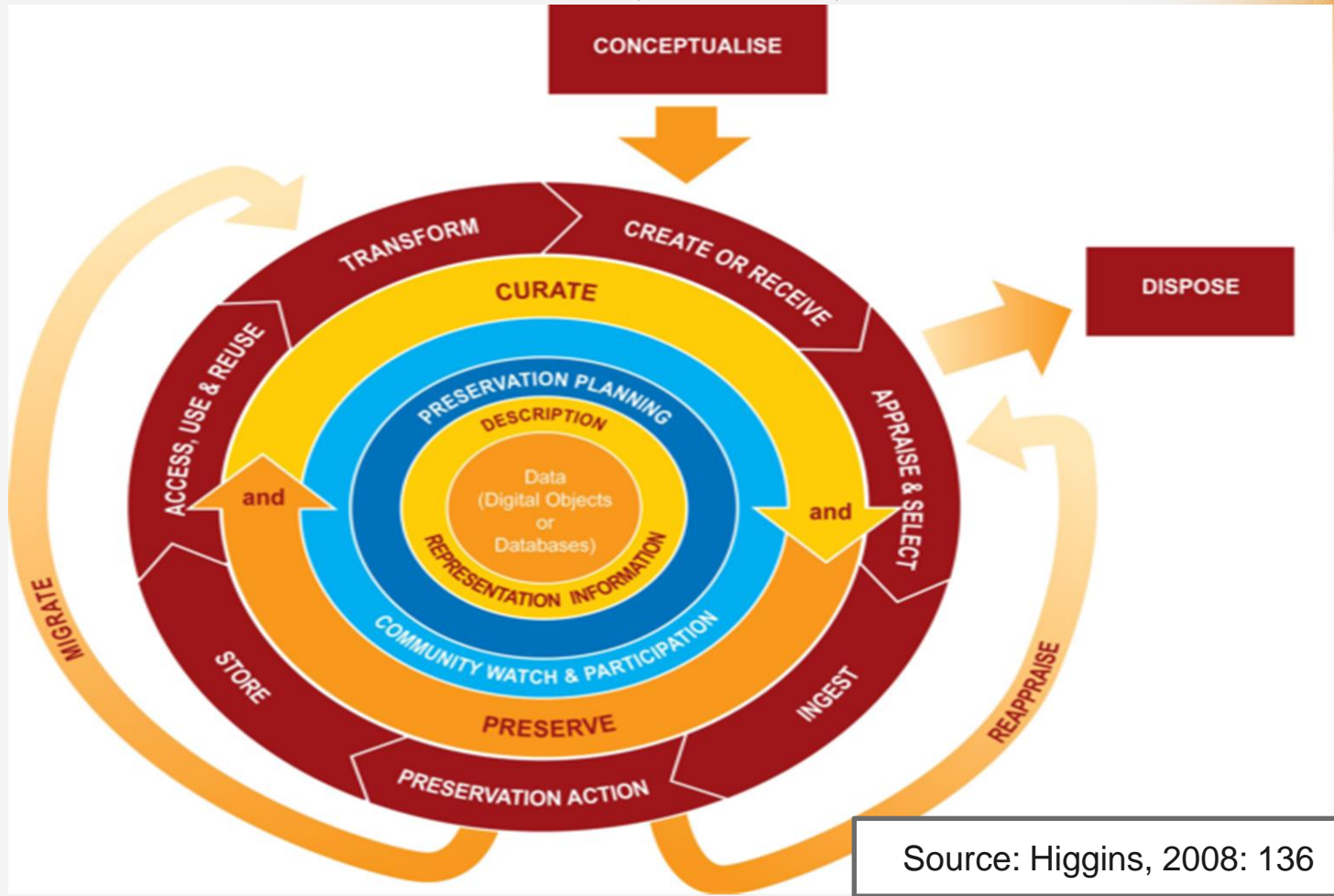
Source: Higgins, 2008: 136

METHODOLOGY

Philosophical Stance
Pragmatism view

Research Approach
Concurrent
Mixed Method

Research Design
Survey design applied
within Multiple case
studies
Interviews

Data Collection
Questionnaires

Study Area
Six Kenya's agricultural
research institutes

Population
- Directors of institutes
- Heads of research,
  Researchers, Heads IT &
  Heads librarian

Sampling
Simple random sampling
(142 researchers)
Census (41- Directors,
HoDs, IT & Librarians)

Kenya
© WorldAtlas.com

Area	Established Findings
Capturing research data - formats	♦ Data-statistical, Database, Images, Spreadsheet, Scanned document, GIS, CAD, Web, Video, Data XML, Audio
Appraisal of research data	♦ Repeatability, Science/historical value, Reuse value, Substantiveness, Complementary/added value, Cost-effectiveness, Access, Uniqueness, Volume
Tools applied in appraisal of data	♦ research institute's policy

Area	Established Findings
Description of research data (metadata)	♦Hard drives, handwritten notes in lab notebooks, column and row labels in spreadsheets, creating descriptive metadata for each dataset, and saving the descriptions with datasets on a hard drive.
Storage and preservation of research data	♦PC hard drive, hard drive of the instruments which generated the data, External drivers, research institutes servers or repositories, D server, CD, Cloud services, USB.
Research data Preservation	♦ >10 years, 5-10 years, 1-5years, <1 year ♦ Majority - unaware

Area	Established Findings
Research data access	♦Group researchers, Other researchers in the research institute, Researchers from external research institutes, the public
Methods of availing of research data	♦External storage device (USB drive, CD/DVD), Hard copy or print, E-mail, Don't share data, Depositing them in an institutional repository, Submitting them to a journal to support publication, Data portal or database-driven website, Collaborative web space (wiki, blog, Google Docs), Deposit them with a specialist data centre

Area	Established Findings
Sharing research data access	♦ External storage devices, hard copy/print, e-mail, institutional repository, and journal publication
Research data reuse reasons	♦ Potential new data use, encouraging scientific enquiry, re-analysis of data can lead to powerful insights, promote innovations, avoid duplication, and reduce the cost of doing research
Responsibility for research data curation	♦ There was no clear unit/department/person responsible for research data curation

- Establish a formal data governance structure

- Enact RDM   legislation

- RDM policies and regulations revision

- Establishment of an RDM department in the library

- Enhance RDM human resource capacity

- Awareness, advocacy, data literacy programmes

- Enhance agricultural research institutes' technical infrastructure

Thank you