INTRODUCTION TO
DATA CURATION

Makaba Macanda
&
Ronell Bezuidenhout
Definitions

• A relatively new discipline with many different definitions

• **Research data**: Research data, unlike other types of information, is collected, observed, or created, for purposes of analysis to produce original research results

http://www.ed.ac.uk/is/data-management
• **Data curation**: the curation of records or measurements of information (“data”). Those scientific measurements or records (“data”) are further distinguished from the computer science meaning of “data” to refer to any type of digitally encoded information.

[http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1027&context=lib_dean](http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1027&context=lib_dean)
Research data management

Is not only:

Data archiving OR

Data backups
Rationale for the Library’s involvement

• “Thus with the experience gained from traditional cataloguing, indexing and organizational skills coupled to those acquired in developing, establishing and maintaining institutional repositories, the time is ripe for academic librarians to explore their role as data curators”
• Data Curation and Libraries: Short-Term Developments, Long-Term Prospects
  http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1027&amp;context=lib_dean

• A new role for academic librarians: data curation
  http://www.era.lib.ed.ac.uk/handle/1842/3207
Data Management concepts, process, ...

- Data Ownership
- Data Collection
- Data Storage
- Data Protection
• Data Retention
• Data Analysis
• Data Sharing
• Data Reporting
UK Digital Curation Centre’s Lifecycle Model

CONCEPTUALISE

CREATE OR RECEIVE

CURATE

DESCRIPTION

PREPARATION PLANNING

Data (Digital Objects or Databases)

PREVENTION ACTION

STORE

REAPPRaise

MIGRATE

ACCESS, USE & REUSE

TRANSFORM

UNISA

university of south africa LIBRARY
Why manage research data?

• Data management is one of the essential areas of responsible conduct of research.
• Before starting a new research project, the researchers and or the research teams must address issues related to data management.
By managing your data you will:

– Meet funding body grant requirements.
– Ensure research integrity and replication.
– Ensure research data and records are accurate, complete, authentic and reliable.
– Increase your research efficiency.
– Enhance data security and minimise the risk of data loss.
– Prevent duplication of effort by enabling others to use your data.
– Comply with practices conducted in industry and commerce.
– Save time and resources in the long run.

http://www.ed.ac.uk/is/data-management
Levels of research data management

- **International** e.g. World Data Centre on Climate

- **National** e.g. Very Large Database Initiative (DST / Meraka, CSIR); NeDICC (Network of Distributed Data & Information Curation Centres) Initiative

- **Campus** e.g. repositories for open access
• Data Management survey at Unisa: March-July 2011

• https://www.surveymonkey.com/s/BGX9L2N
Data Curation Proposal

• Unisa departments involved
• National initiatives
• Short term management of data
• Description of data to be archived
• Acquiring of data
• Processing of data and system description
• Metadata content and format
• Quality assurance and quality control
• Training and support
• Budget/Funding
Conclusion
The end, for now