Reflections on the RDM position in South Africa
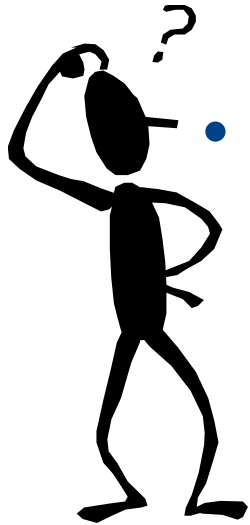
Developing Research Data Management Services
LIASA WCHELIG/HELIG/DCC Workshop
Cape Town
27 March 2014

Lucia Lötter

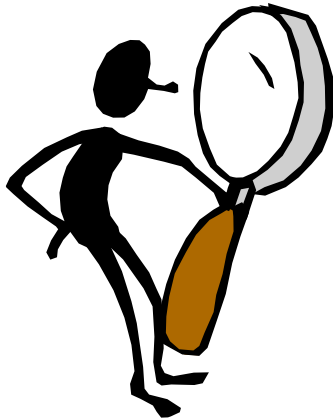Research Methodology Centre

Objectives of this presentation

- A mapping of the current landscape of entities sharing an interest in RDM in SA

- Organisational responses to the RDM challenge
  - Data producers
  - Data centres / libraries
  - National initiatives

- Reflection on the status of RDM in SA
  - Achievements
  - Challenges
  - Opportunities
  - Suggestions & encouragement

HSRC
Human Sciences
Research Council

Reflecting on RDM in SA

- RDM complexity
  - Research data (small vs big)
  - Research data management (activities, levels)
  - Enabling services
  - Roles

- A small and specific lens
  - Specific focus
  - Not exhaustive (organisations, activities, organisational responses)
  - Transparency of information

The RDM landscape in SA

Entities

- Higher Education Institutions (HEIs)
- Research councils
- NRF (National Research Facilities)
- Government agencies
  - NSO – Stats SA, DST, NRF, DHET, other departments, NIPMO, other national entities (SANPARKS, SANBI)
- NGOs
- Private sector
- Data centres/libraries (projects, funders, themes, disciplines)

Roles

- Data producers
- CoPs & Associations
- Funders
- Service providers (IT Infrastructure)
- Distributors
- Policy makers
- Capacity building
- Data users

HSRC
Human Sciences
Research Council

Responses to the RDM challenge
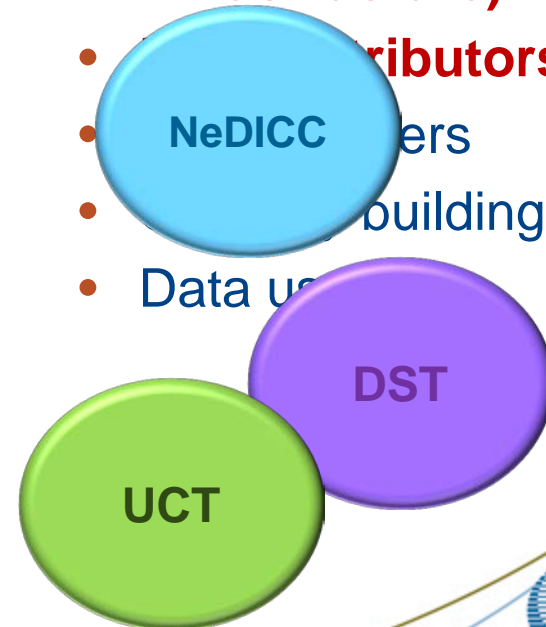## Various initiatives

Entities

- Higher Education Institutions (HEIs) - Universities
- Research councils
- NRF (National Research Facilities, Centres of Excellence)
- **Government agencies**
  - NSO – Stats SA, **DST, NRF**, DHET, other departments, NIPMO, national ~
- NGOs
- Private sector
- **Data centres/libraries** (projects, funders, themes, disciplines)

Roles

- Data producers
- **CoPs** & Associations
- Funders
- **Service providers (IT Infrastructure)**
- ~ributors
- ~ers
- ~building
- Data us~

NRF

NeDICC

DST

UCT

HSRC
Human Sciences
Research Council

Various initiatives

- International representation
  - Country membership to ICSU, CODATA [NRF]
- Promotion of sound RDM practices
  - OECD signatory on publicly funded research [DST]
  - RDA [Individual membership]
- CoP [NeDICC: ARC, ASSAF, CSIR, HSRC, NRF, UNISA, UP, WITS]
- National Integrated Cyber-Infrastructure (NICI) [DST]
  - High Performance Computing (CHPC-Rosebank)
  - South African National Research Network (SANReN)
  - South African GRID (SA-GRID)
  - Data Intensive Research Infrastructure for SA (DIRISA)

HSRC
Human Sciences
Research Council
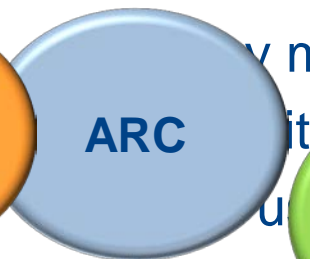
Responses to the RDM challenge
Selected data producers

Entities

- Higher Education Institutions (HEIs) - **Universities**
- **Research councils**
- NRF (**National Research Facilities**)
- Government ~~~~
  - ~~~~SO – Stat~~~~
  - ~~~~T, other depart~~~~
  - ~~~~MO, national ent~~~~
- ~~NGOs~~
- Private se~~~~
- Data centres ~~~~es (proje~~~~ funders, themes, disciplines)

Roles

- **Data producers**
- CoPs & Associations
- Fund~~~~
- Serv~~~~ (IT Infrast~~~~
- Data distributors
  - ~~~~y makers
  - ~~~~ity~~~~
  - ~~~~u~~~~

UP

WITS

CSIR

UNISA

NRF

ARC

US

UCT

HSRC

Selected data producers

❶ **Do we know?**

❷ **What are the external requirements?**

❹ **Are responses formalised?**

RDM Services

Are training, support and guidance provided to researchers & students? ❸

❺ **Cater for costs associated with RDM?**

❻ **Are platforms and services for the storage, preservation & sharing in place?**

❼ **Do we have skilled staff?**

Based on DCC Mini Cardio Quiz (http://cardio.dcc.ac.uk/quiz/)

Do we know?

- Situation analysis, determined need and readiness with survey of researches [CSIR, UCT, UNISA, UP, WITS], Pilot survey amongst departments to feed into institutional needs analysis with aim of articulation of the RDM challenge in a business case [SU]

- Engaging with role-players around an institutional approach to RDM [SU], Plan to establish whether there are institutional initiatives for proper, corporate wide data management and curation [ARC]

- High-level business requirements on RDM [UNISA]

- Importance of proper data management are generally accepted. [CSIR, HSRC, NRF, UNISA, UCT, UP]

- Engagement with management [CSIR, SU, UP, WITS], directive from management mandated to investigate RDM [UNISA], mandate from management [HSRC,UCT], mandated by act and entrenched as a strategic objective [HSRC], process to gain management buy-in is planned [ARC]

- Internal advocacy [HSRC, UP]

- RDM prototypes / case studies (RDM as integral part of the research process; not a separate activity) [UP]

- Conferences, workshops (local & abroad)

- Self-learning and networking

What are the external requirements?

❷

- Study funder requirements and relevant legislation. Engagement with NIPMO (National Intellectual Property Management Office) [HSRC]

- Request information about requirements as part of a data preservation and sharing plan for each research project [HSRC]

- Provide assistance to researchers in planning for RDM to meet the requirements of funders [HSRC]

- Legislative and funder requirements considered in policies and procedures and research contracts. [HSRC]

HSRC
Human Sciences
Research Council

Are responses formalised?

- RDM policies in varying stages of development [CSIR, HSRC, UCT, UNISA, UP, WITS]

- Policies of National Research Facilities, e.g. SAEON [NRF]

- UP has an outdated RDM policy (2007) – a new one will be developed this year, in collaboration with other units, departments and researchers [UP]

- Ethics policy reviewed to incorporate RDM [UNISA] , review [HSRC]

- Data sharing policy, as well as SOP for the curation of data in place. Data preservation policy submitted to management for approval. Quarterly and annual report statistics (audited as performance against annual targets) [HSRC]

- Link RDM to records management policy [CSIR]

- High-level report on RDM internationally, nationally and at UP [UP]

- High-level business requirements on RDM [UNISA]

HSRC
Human Sciences
Research Council

Are training, support & guidance provided?

④

- Engagement with researchers and students about RDM, RDM training provided by faculties and library [WITS]

- Planned capacity building, advocacy, training [UCT, UNISA],

- Interaction with researchers at National Research Facilities, e.g. SAIAB, SAEON [NRF]

- Researcher training around data management and documentation, data deposit, secondary data use [HSRC]

- Engagement with researchers [UP Case studies, UNISA]

- Review of data preservation and sharing plans [HSRC]

- One-on-one consultation with researchers around data deposit [HSRC]

HSRC
Human Sciences
Research Council

Are platforms and services in place?
(Storage, preservation & sharing)

- Project based implementation of best practice RDM [CSIR, NRF, UNISA, UP, WITS]

- Library is proposed as repository of last resort [WITS]

- Big data – aware that compliance to international RDM standards will be required, Sharing of data with large data repositories – discipline specific, WDS. [Most universities, CSIR, NRF National Research Facilities]

- Institutional repositories which can be extended to house data [most organisations]

- Link data sets to accompanying theses/dissertations [UCT, SU]

- Preservation of content hosted in digital repository - backups, use of dSpace for preservation [SU]

- Data shared on demand [ARC, CSIR, NRF National Research Facilities]

- Platform for data curation (metadata and file repository), including sharing with rule based access management to metadata and files, long term preservation to be outsourced [HSRC]

- System requirements and specifications, trail period evaluation of software [UNISA] and proof of concept implementation [UP]

- Data centres / libraries [DataFirst, SADA, Stats SA, WDS, SAEON]

HSRC
Human Sciences
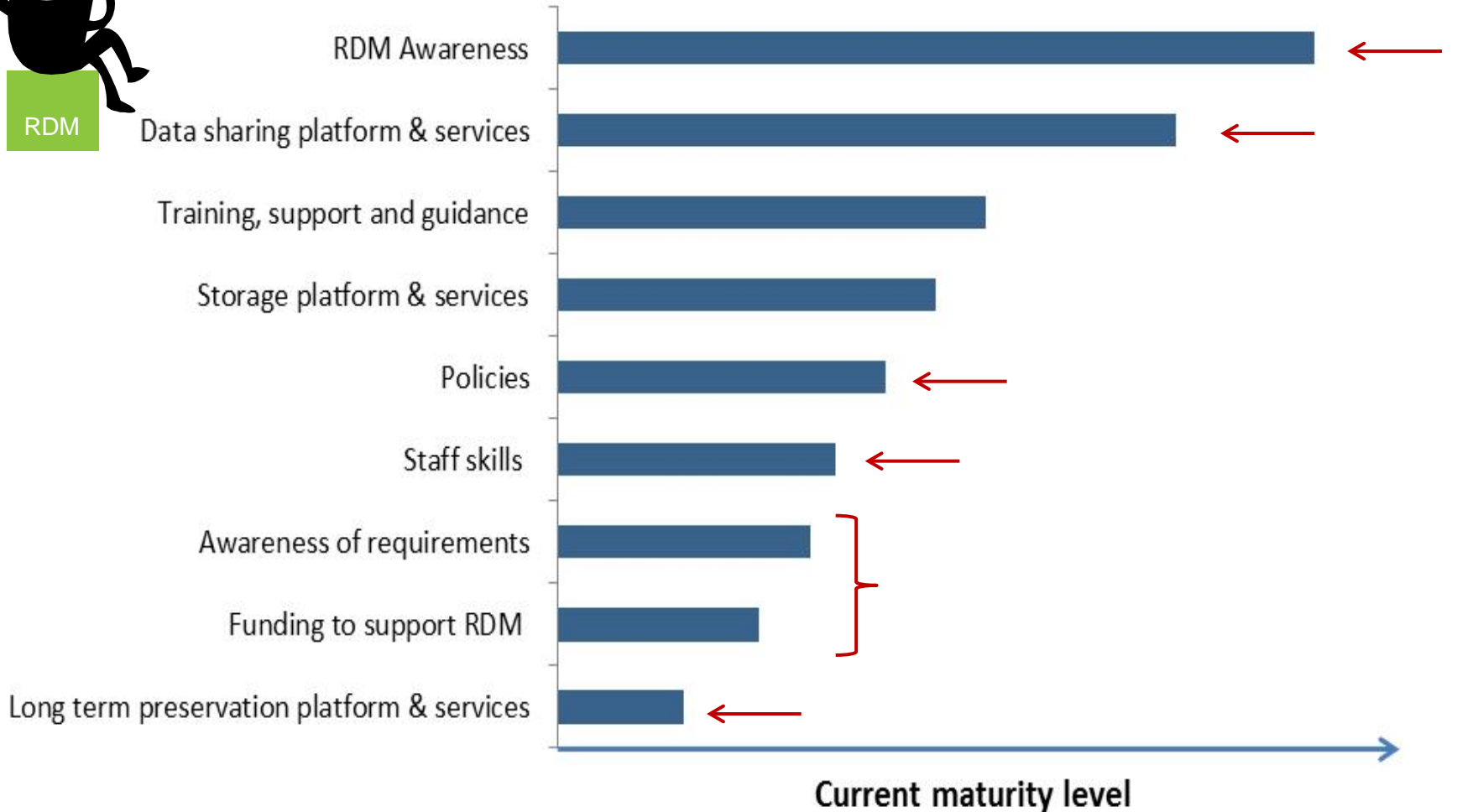Research Council

Funding to support RDM?

- Annual overhead budget for infrastructure, general services and operational cost supplemented by cost recovery from research programmes on which behalf data is curated. [HSRC]

- Project based funding [CSIR - short term and not yet fully transparent]

- Operational budget [UCT]

- Budgets for institutional repositories
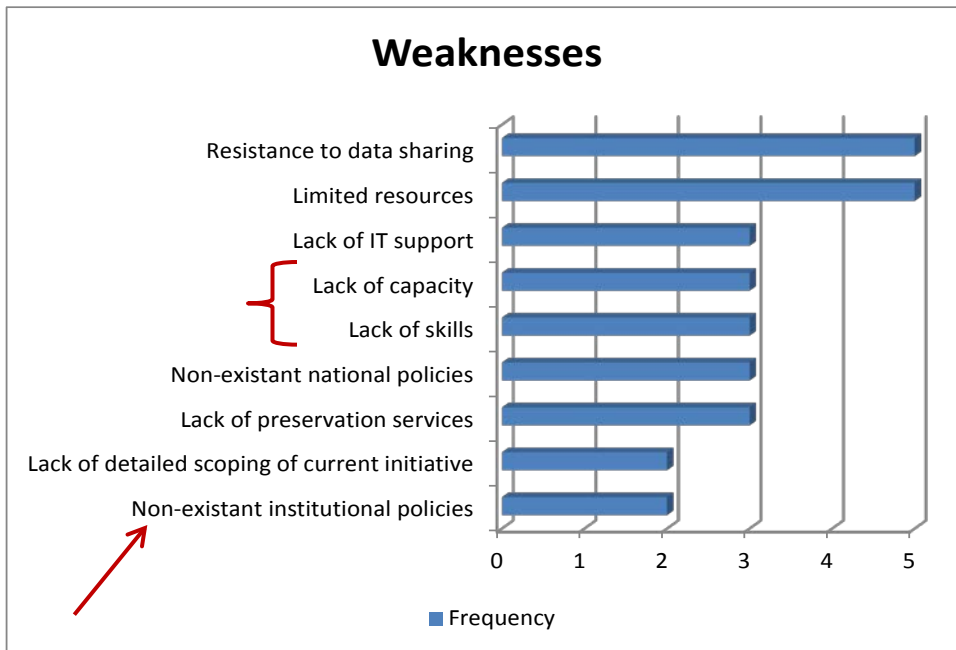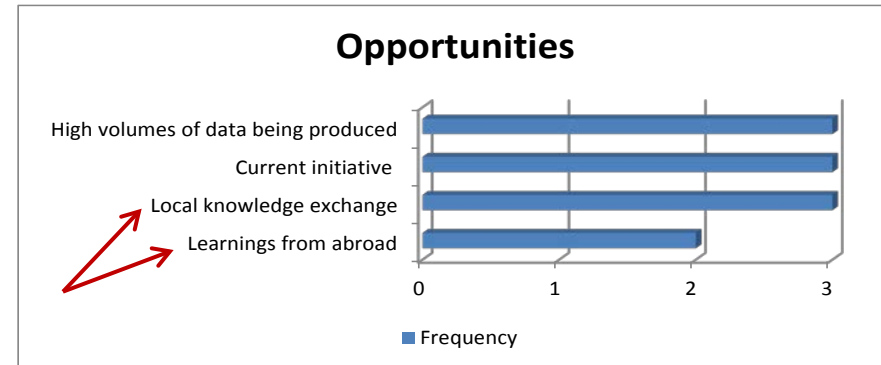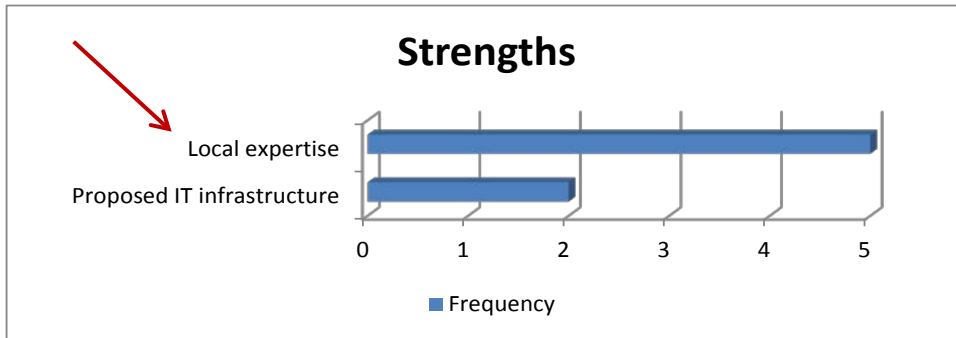
Do we have skilled staff?

- Appointment of staff with RDM as part of KPAs. [CSIR, HSRC, NRF National Facilities, UCT, UNISA, UP, WITS]

- Capacity building through NeDICC (CoP, ad hoc training workshops)

- In-house training and mentorship of Research Data Curators [HSRC]

- Opportunities to attend conferences, courses and fact finding abroad

- Knowledge sharing [HSRC]

- Conferences, workshops (local & abroad)
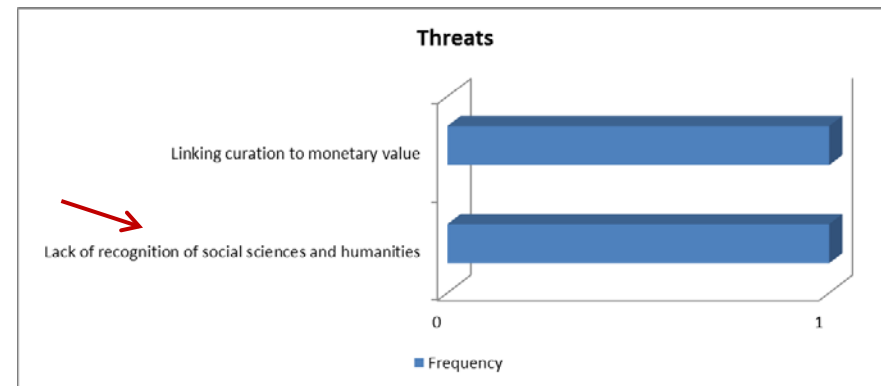
- Self-learning and networking

HSRC
Human Sciences
Research Council

Perception of maturity of RDM responses
Maturity relative to others dimensions

RDM

RDM Awareness	
Data sharing platform & services	
Training, support and guidance	
Storage platform & services	
Policies	
Staff skills	
Awareness of requirements	
Funding to support RDM	
Long term preservation platform & services	

Current maturity level

Analysis of current status

Strengths

Opportunities

Weaknesses

2014 ?

Threats

SWOT Analysis (Preserving and providing access to South African social science and humanities research data: 5 November 2012)

The RDM challenge

CARDIO
collaborative assessment of research data infrastructure and objectives

- Collection development
- Depositor support
- Digital object management
- Promotion of secondary data discovery and use

- Assessment
- Planning
- Bench marking

- Policies
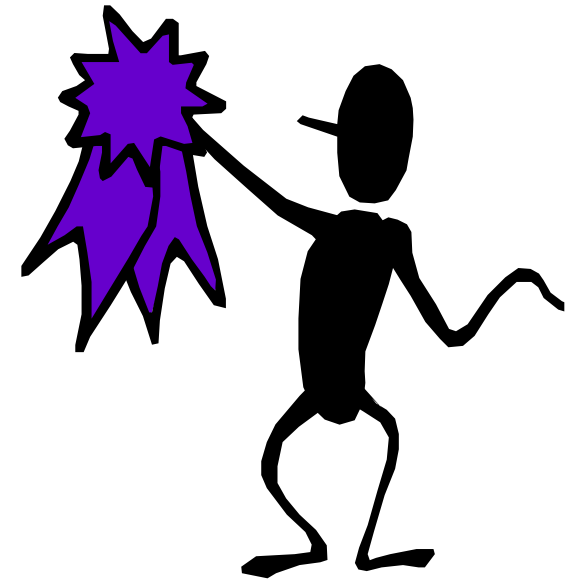- Procedures
- Technology
- Capacity
- Financial resources

Models, standards
- OAIS
- TRAC

HSRC
Human Sciences
Research Council

What can we do?

- Lobby for national policies and support

- Continued advocacy in organisations

- Mutual assistance

  - Capacity building

  - Sharing of experiences, knowledge

  - Support shared goals
                                          unique
  - Encourage each other – celebrate achievements ☺ and ☺
                                          ∧
- Learn

- Capitalise on momentum

- Press on

HSRC
Human Sciences
Research Council

HSRC
Human Sciences
Research Council

Changing lives of
ordinary people
through human and
social sciences

Technology

LIFE ORIENTATION

Mathematics

Thank you

llotter@hsrc.ac.za

www.hsrc.ac.za

Organisation list

- Government entities
    - DHET - Higher Education and Training
    - DST – Department of Science and Technology
    - NIPMO - National Intellectual Property Management Office
    - NRF – National Research Foundation
    - SANPARKS – South African National Parks
    - SANBI – South African National Biodiversity Institute
- Universities
    - SU – Stellenbosch University
    - UCT – University of Cape Town
    - UNISA – University of South Africa
    - UP – University of Pretoria
    - WITS – University of the Witwatersrand
- Science councils
    - ARC – Agricultural Research Council
    - CSIR – Council for Scientific and Industrial Research
    - HSRC – Human Sciences Research Council
- National Research Facilities
    - HartRAO - Hartebeesthoek Radio Astronomy Observatory
    - iThemba LABS - Laboratory for Accelerator Based Sciences
    - NZG - National Zoological Gardens of South Africa
    - SAAO - The South African Astronomical Observatory
    - SAEON - South African Environmental Observation Network
    - SAIAB - South African Institute for Aquatic Biodiversity