

# Reflections on the RDM position in South Africa

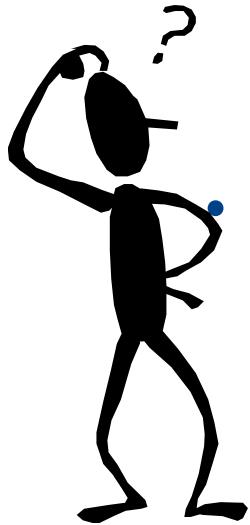
**NeDICC / DCC Workshop  
Pretoria  
11 August 2015**

Developed by  
Lucia Lötter

Presented by  
Elias Makonko

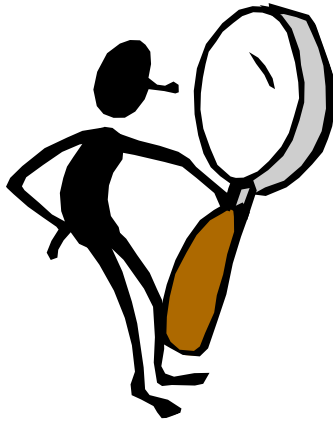
# Objectives of this presentation

- A mapping of the current landscape of entities sharing an interest in Research Data Management (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



# 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
  - Lack of shared understanding

# Responses to the RDM challenge

## Role players

### 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 entities
- NGOs
- Private sector
- Data centres/libraries (projects, funders, themes, disciplines)

### Roles

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

# Various initiatives

- International representation
  - Country membership to ICSU, CODATA [NRF]
- Promotion of sound RDM practices
  - OECD signatory on publicly funded research [DST]
  - Research Data Alliance (RDA) [Individual membership]
  - Open access statement, 2015 [NRF]
- CoP [NeDICC: ARC, ASSAF, CSIR, HSRC, NRF, UNISA, UP, WITS]
- Universities [CPUT, UJ, SU]
- National Integrated Cyber-Infrastructure (NICIS) [DST]
  - High Performance Computing (CHPC)
  - South African National Research Network (SANReN)
  - South African GRID (SA-GRID)
  - Data Intensive Research Infrastructure for SA (DIRISA)
- Collaborative capacity building initiatives (DCC & SA)

# Reflecting on the RDM challenge in terms of..

**1** Do we know?

What are the external requirements?

**2**

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



Are responses formalised?

**4**

**3**



RDM  
Services

**5**

Cater for costs associated with RDM?

**6**

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

**7**

Do we have skilled staff?

# Do we know?

- Situation analysis, determined need and readiness with survey of researchers [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], requirements gathering and analysis through pilot project with research centres for the development of eResearch infrastructure [CPUT]
- Engaging with role-players around an institutional approach to RDM [SU, UP], Plan to establish whether there are institutional initiatives for proper, corporate wide data management and curation [ARC], Formed an institutional working group with members cutting across the institution for policy and eResearch service development [CPUT]
- Working on implementation plan for RDM (framework and architecture) [UP]
- High-level business requirements on RDM [UNISA]
- Importance of proper data management is generally accepted. [CSIR, HSRC, NRF, UNISA, UCT, UP, CPUT]

# 1 Do we know? (cont.)

- On a university wide level RDM is seen as important, nothing has been acknowledged at Faculty level [UJ]
- Engagement with management [CSIR, SU, UP, WITS, CPUT], 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, CPUT]
- RDM prototypes / case studies (RDM as integral part of the research process; not a separate activity) [UP, CPUT]
- Conferences, workshops (local & abroad)
- Self-learning and networking



# What are the external requirements?

- 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, CPUT]
- Provide assistance to researchers in planning for RDM to meet the requirements of funders [HSRC, CPUT]
- Legislative and funder requirements considered in policies and procedures and research contracts. [HSRC]
- NRF Statement on open access to publicly funded research outputs – as of April 2015

# Are responses formalised?

- RDM policies in varying stages of development [CSIR, UCT, UNISA, UP, WITS ]
- No policies or processes [UJ]
- Policy approved by university council in July 2014 [CPUT]
- Policies of National Research Facilities, e.g. SAEON [NRF]
- UP has an outdated RDM policy (2007) – a new one is in development, 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 and sharing policies approved. Quarterly and annual curation 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]
- High-level report on RDM internationally, nationally and at CPUT [CPUT]

# Are training, support & guidance provided?

- Engagement with researchers and students about RDM, RDM training provided by faculties and library [WITS]
- Engagement with research centres and units on RDM [CPUT]
- Planned capacity building, advocacy, training [UCT, UNISA, CPUT],
- 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, CSIR, CPUT]
- Review of data preservation and sharing plans [HSRC]
- One-on-one consultation with researchers around data deposit [HSRC, CPUT]
- Post-graduate training in Digital Curation [UCT's Library and Information Studies Centre (LISC)]
- Workshops [NeDICC, UWC/CHELSEA]
- eResearch Africa Conference [ASAUDIT]

# 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]
- DIRISA as repository of last resort
- Library coordinates the development of an eResearch and Communication Infrastructure (eRIC), further developing Mediatum to support research lifecycle [CPUT]
- 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]

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

- 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, trial period evaluation of software [UNISA] and proof of concept implementation [UP]
- Data centres / libraries [DataFirst, SADA, Stats SA, WDS, SAEON]

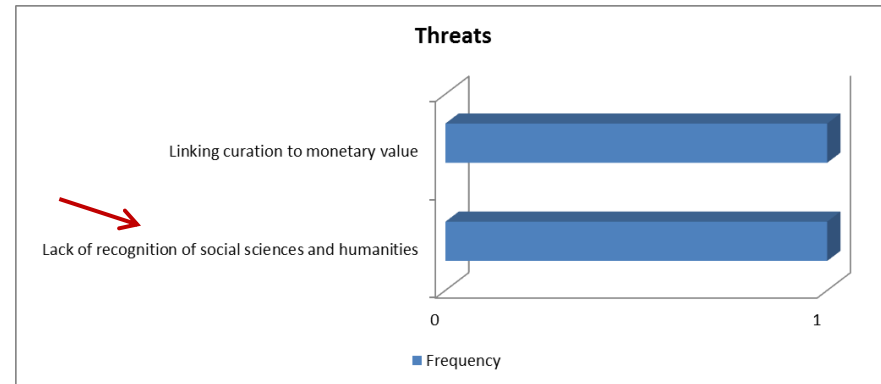
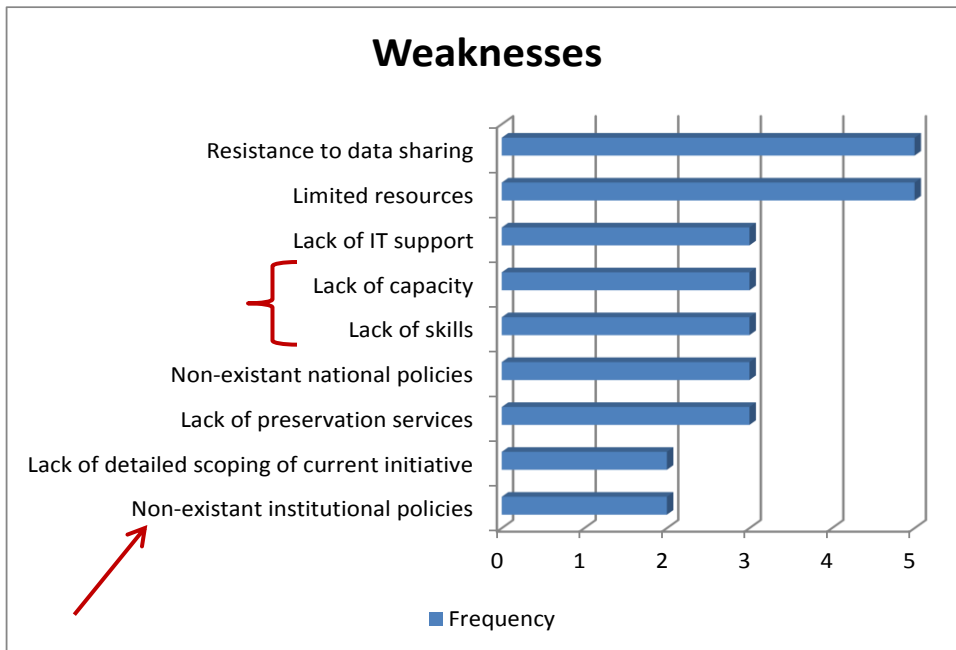
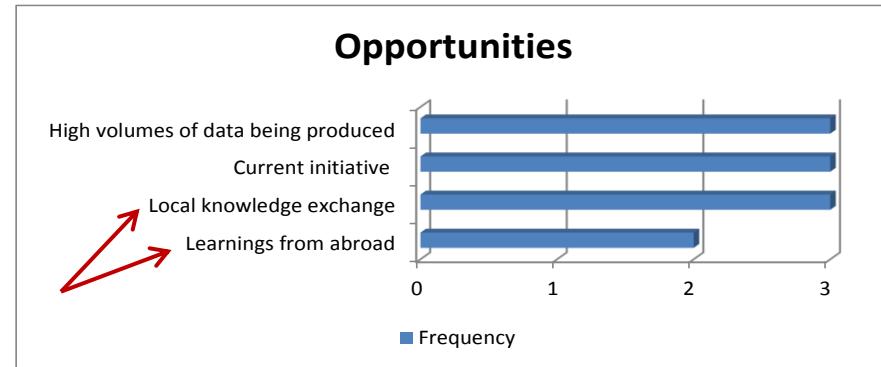
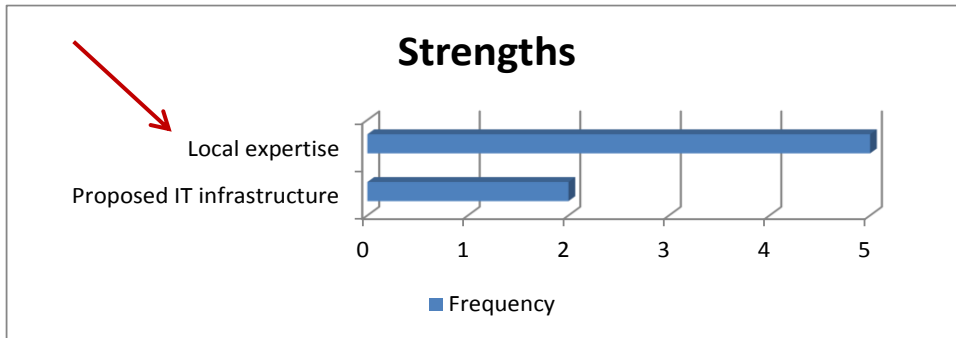
# Funding to support RDM?

- Annual overhead budget for infrastructure, general services and operational cost supplemented by cost recovery from research programmes on which behalf data is curated [HSRC]
- Funding as part of research grants [HSRC]
- Project based funding [CSIR - short term and not yet fully transparent]
- Operational budget [UCT]
- Operating within the Library IT budget [CPUT]
- Budgets for institutional repositories
- No consideration of funding (on a departmental and researcher level) [UJ]

# Do we have skilled staff?

- Appointment of staff with RDM as part of KPAs. [CSIR, HSRC, NRF National Facilities, UCT, UNISA, UP, WITS]
- Appointment of systems developer (programmer) for eResearch project [CPUT]
- 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) [DCC LIASA / NeDICC, UWC / Chelsa]
- Presentations (UP)
- Self-learning and networking
- Focus on developing researchers and not the management of the research data [UJ]

# 2012 SWOT Analysis





# 2014

## (Librarians' views on the changing research data management landscape)

- Awareness of RDM
  - Policy and framework situation for RDM is haphazard and very much in its infancy.
  - A need for improved communication between relevant government agencies, institutions and researchers.
  - Majority of respondents are at least aware of the absence of institutional RDM policy and the importance of this absence.
- Policy, policy-making and responsibilities
  - Policy-making should be a collaborative process.
  - Coordination: Collaboration between government and institutional consortia vs led at national level by government (Nearly 50/50).
  - Curate data: Professional / disciplinary body / ICT, not library.

# 2014 (cont.)

- Skills and resource gaps
  - IT skills
  - Guidance and support, curation skills and training, policy frameworks, curation tools and software, storage and network
- Sharing and Open Data
  - Researchers at their institutions are receptive to the idea of sharing – But not mine?

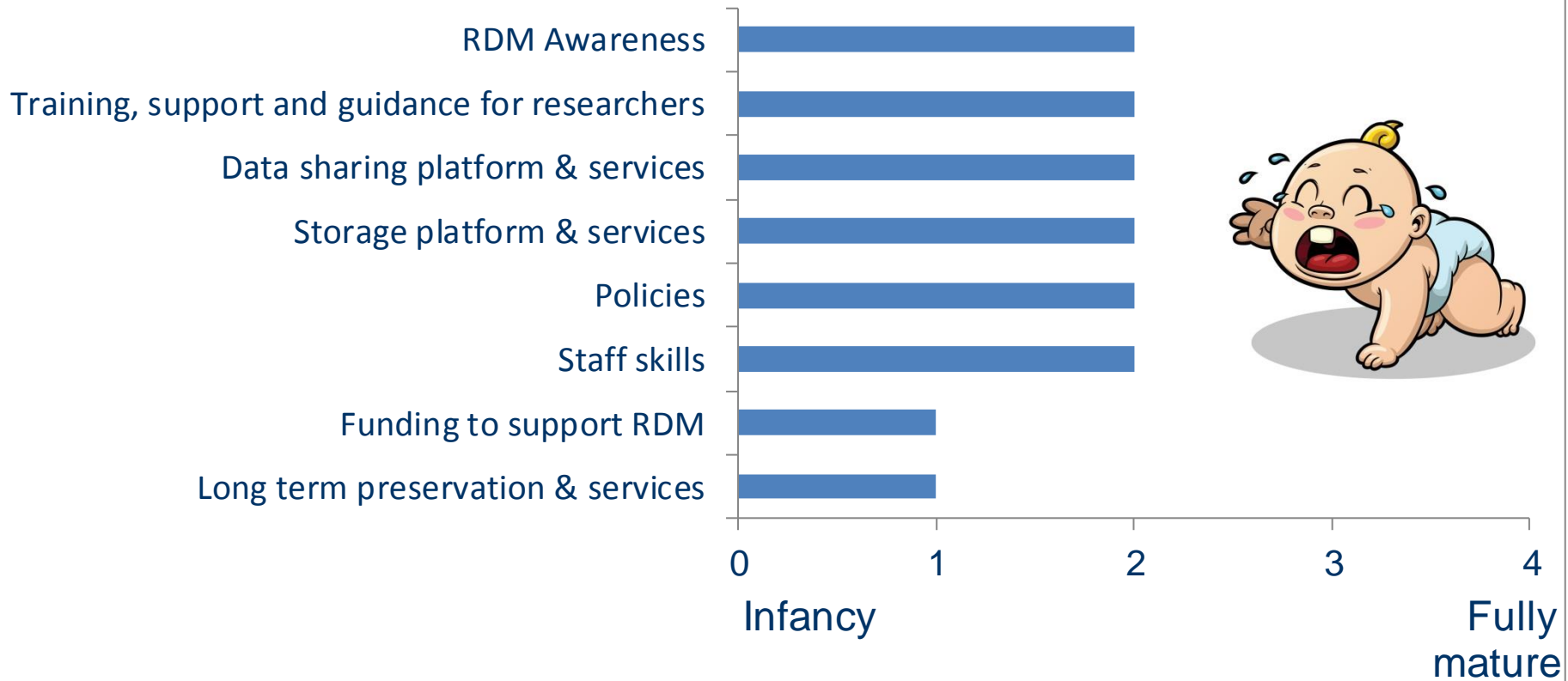
Kahn, M. Higgs, R. Davidson, & J. Jones, S. 2014 Research Data Management in South Africa: How We Shape Up. Australian Academic & Research Libraries, 45, 4, 296-308.

## 2015 ?

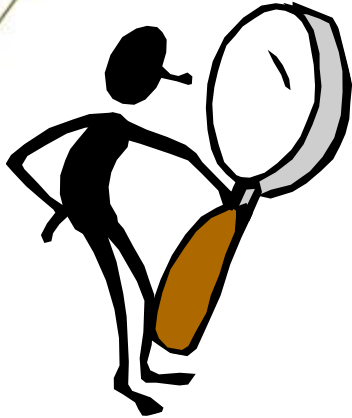
# Perception of maturity of RDM responses

## Maturity relative to other dimensions

### Median maturity level perceptions of NeDICC members (2015)



# Do we really know?



- Contradictions
- Roles and responsibilities not clear (top down vs bottom up?)
- High expectations
- Perceptions vs reality
- Perceptions of own organisation, known others vs the country as a whole
- Whose perceptions / reality?
  - Are we asking the right people / organisations?
  - Who are the people / characteristics of the respondents targeted?
  - What is the RDM knowledge / skills base / approach of the sample?
  - Methodology?

# The RDM challenge

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



- Assessment
- Planning
- Bench marking



- Policies
- Procedures
- Technology
- Capacity
- Financial resources



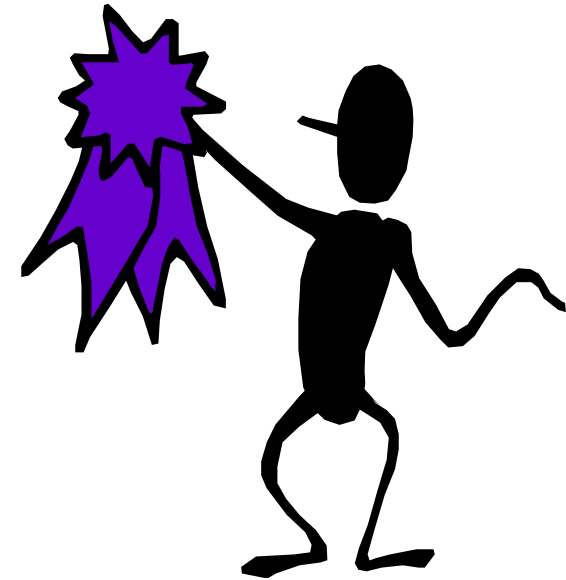
Models, standards

- OAIS
- TRAC

# 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
  - Encourage each other – celebrate achievements
- Learn
- Capitalise on momentum
- Press on

unique



# Acronyms

- 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
  - ASSAf - Academy of Science of South Africa
  - NSO – National Statistics Office
  - Stats SA – Statistics South Africa
  - SADA – South African Data Archive
- Universities
  - CPUT – Cape Peninsula University of Technology
  - SU – Stellenbosch University
  - UCT – University of Cape Town
  - UNISA – University of South Africa
  - UP – University of Pretoria
  - WITS – University of the Witwatersrand
  - UJ – University of Johannesburg
  - UWC – University of the Western Cape
- Associations
  - CHELSA – Committee of Higher Education Libraries of South African
  - ASAUDIT – Association of South African Directors of Information Technology
  - LIASA – Library and Information Association of South Africa

# Acronyms

- International entities
  - ICSU - International Council for Science
  - CODATA - Committee on Data for Science and Technology
  - OECD - Organisation for Economic Co-operation and Development
  - DCC – Digital Curation Centre
  - WDS – World Data System
- 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



**Changing lives of  
ordinary people  
through human and  
social sciences**



**Thank you**

[llover@hsrc.ac.za](mailto:llover@hsrc.ac.za)

[www.hsrc.ac.za](http://www.hsrc.ac.za)