

DATA CURATION: A MEANS TOWARDS AN END

Dr Lucia Lötter

Human Sciences Research Council, Pretoria, South-Africa, llotter@hsrc.ac.za



HUMAN SCIENCES RESEARCH COUNCIL

THE HSRC'S FOCUS ON RESEARCH DATA

The HSRC is a statutory research organisation that advances social sciences and humanities for public use. It undertakes research projects addressing challenges relating to **poverty, economic growth and job creation, quality of education, skills development, service delivery, crime, youth, families and social cohesion, health and wellbeing, HIV/AIDS and STI control, nutrition and food security, as well as the national system of innovation.** The focus on social science and humanities data has been articulated as an objective in the HSRC Act 17 of 2008 [1]. In response to this requirement the HSRC committed to the implementation of data curation as an organisational objective [2]. Research data is shared in order to make evidence available which speaks to the challenges of society for the purposes of **research, decision making, public dialogue, as well as monitoring and evaluation.**

The HSRC intends to make its research data accessible and to ensure its future survival and usability by **preserving and disseminating** data sets for secondary use to the maximum extent possible [3]. The HSRC adopted the stance that **curation**

embraces and goes beyond that of enhanced present-day re-use and of archival responsibility, adding value through the provision of context and linkage by publishing data in ways that facilitates re-use and promotes accountability and integration as articulated as a vision for digital curation [4].

DATA CURATION AS A PREREQUISITE FOR DATA SHARING

For data to be shared, it has to be **usable**, i.e.

- understandable,
- of good quality,
- readily available and
- accessible.

The implementation of data curation can only happen within a particular **context** in terms of the

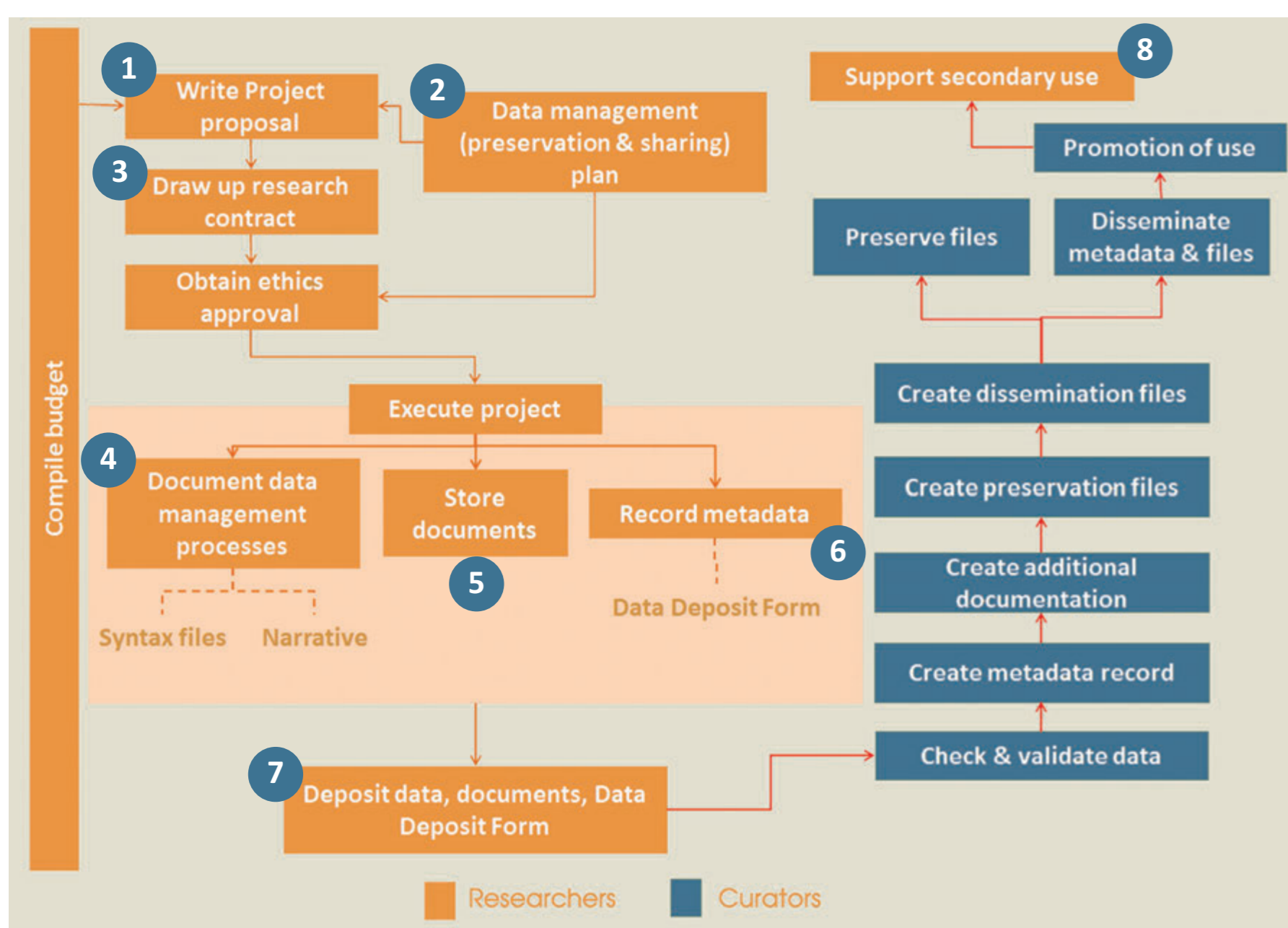
- nature of the data,
- responsible conduct of research,
- legal obligations and
- organisational culture and commitment.

Data curation activities are aimed at delivering on these requirements.

THE IMPLEMENTATION OF RESEARCH DATA CURATION AT THE HSRC

The implementation at the HSRC is an ongoing process and consists of numerous activities around data and process flows, implementing suitable technologies, developing policies and procedures, as well as capacity building and change management.

Process flow

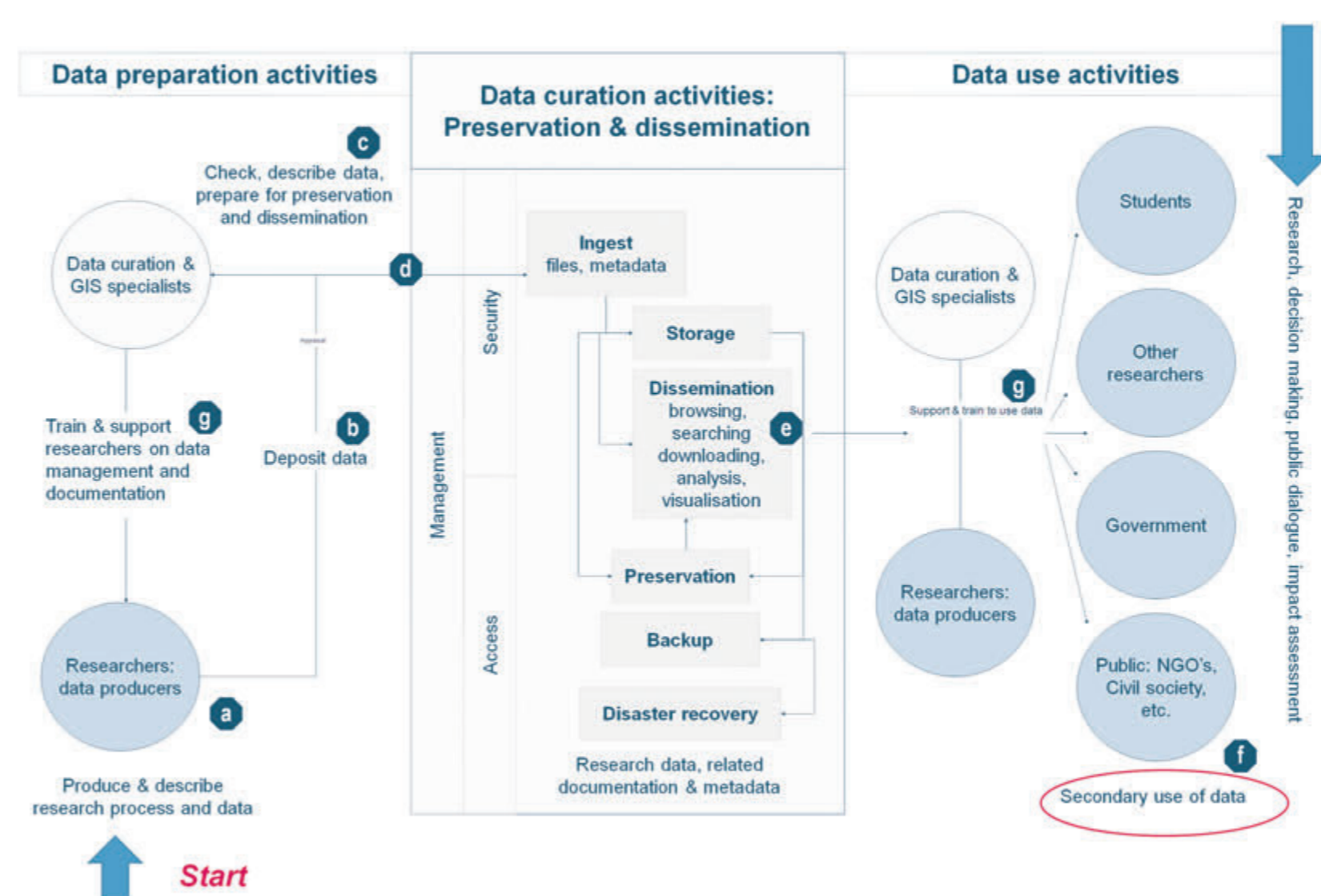


The high level **process flow** within which data curation activities take place, focuses on the following:

- Including preservation and dissemination issues in data management plans of research proposals
- Dealing with data preservation and sharing matters as part of the HSRC's ethics review process
- Addressing issues of ownership and secondary use in research contracts
- Promoting proper data management and documentation throughout the execution of research projects
- Depositing of data and related documents for curation
- Disseminating data and documents
 - Check and validate data (error detection, anonymisation)
 - Create data documentation (create catalogue record in metadata repository, enhance internal metadata (variable and value labels, develop contextual information to facilitate re-use, catalogue related documentation, e.g. questionnaires, code books)
 - Link research outputs to data
 - Create and store (in file repository) dissemination formats and implement access management parameters for metadata and files.
- Preserving data and documents
 - Create preservation formats
 - Storage of preservation files and storage management (multiple copies, multi-media), backup, disaster recovery and security
- Promoting and facilitating secondary use of data

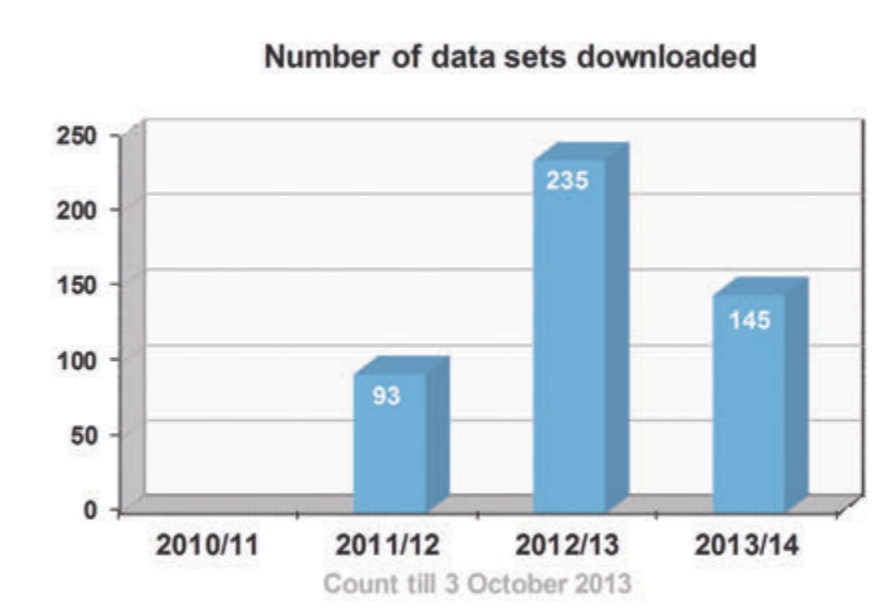
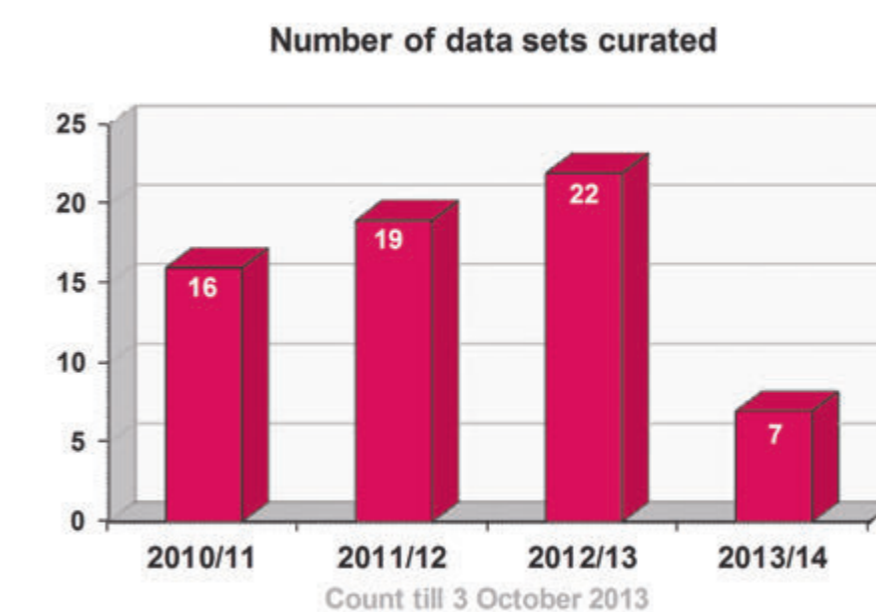
Data curation - A means towards an end

With the end in mind ...



Dissemination on the HSRC's website (<http://www.hsrc.ac.za>)

TO DATE



REFLECTION

Challenges

Lack of a sharing culture and resistance from researchers

Data at high risk of loss, damage or becoming unusable over time

Lack of off-the-shelf technologies

Lack of trained research data curators

Suggested interventions

- ➔ Change management
- ➔ Implement managed access
- ➔ Demonstrate success and build trust
- ➔ Implement policies and procedures around data management, and specifically in terms of data preservation
- ➔ Data management and documentation training
- ➔ Develop bespoke systems
- ➔ In-house research data curation training

REFERENCES

1. SOUTH AFRICA, 2008. Human Sciences Research Council Act 2008. Cape Town.
2. Human Sciences Research Council 2013. Human Sciences Research Council Strategic Plan For The Fiscal Years 2012 - 2017. Pretoria South Africa: Human Sciences Research Council.
3. Human Sciences Research Council 2013. Information Policies and Procedures, HSRC Data Sharing Policy. Pretoria South Africa: Human Sciences Research Council.
4. IEEE. The Digital Curation Centre: A vision for digital curation, 2005.