World Digital Preservation Day
7 November 2019

DIGITAL LIBRARY SERVICES

Approaching Digital Preservation at Scale: A Pilot Programme at UCT Libraries

University of KwaZulu-Natal Special Collections Preservation Conservation Conference 2019

Disaster Prevention Preparedness, Response & Recovery of Collective Collections and E-collections

Thursday, 27th September 2019

UKZN Pietermaritzburg Campus, Pietermaritzburg

UCT Libraries
Digital Library Services

Niklas Zimmer (manager)

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Zimmer, N. (2019):
Approaching Digital Preservation at Scale: A Pilot Programme at UCT Libraries.
Presentation. https://doi.org/10.25375/uct.8982452

CC BY

Who's talking?

DIGITAL LIBRARY SERVICES

What do we mean by Digital Preservation (DP)?

- Digital Preservation is concerned with providing long-term access to digital objects, preserving continuity in form as well as functionality.
- It is not simply a backup of data, because long-term digital preservation must consider format, software and hardware obsolescence, among other issues.
- Although it is possible for anyone to read a page from a book written 100 years ago, the same is not true of (e.g.) a floppy disk containing WordPerfect files from twenty years ago.

Source: Preservica: **A Guide to Making the Case for Digital Preservation**. Presentation. 2014. (Online), Available: https://preservica.com/uploads/legacy/2014/04/A-Guide-to-Making-the-Business-Case-for-Digital-Preservation-2014.pdf

Outline

- **Drivers for Digital Preservation**
  - Business Owners; Value Propositions; Policies & Frameworks
- **Practicalities of Digital Preservation**
  - Best Practices; Tools & Systems; Vendors; Roles; Activities
- **A Data Life-cycle approach to Digital Preservation**
  - Plan & Design
  - Collect & Capture
  - Collaborate & Analyse
  - Discover, Reuse & Cite
  - Share & Publish
  - Manage, Store, Preserve
- **Digital Preservation: an all-inclusive conversation**
  - DPC | World Digital Preservation Day (07.11.2019)
  - Australasia Preserves | A digital preservation CoP
  - NeDICC | Network of Data and Information Curation Communities
  - RDM at UCT Slack Workspace

World Digital Preservation Day
7 November 2019

DIGITAL LIBRARY SERVICES

Drivers for Digital Preservation

- *Business Owners*
- *Value Propositions*
- *Policies & Frameworks*

Business owners

'[Digital Preservation] forms part of the integrated management of digital content and assets in such infrastructures [...]. *In a business environment therefore, DP is the responsibility of data and asset management or legal compliance. In other words, it is driven by IT and legal departments, and is not part of the corporate mission as is the case for Memory Institutions (MIs). The common feature between MIs and Business Corporate Enterprises (BCEs) is the record-keeping aspect. BCEs are safekeeping their organization's records under the broad definitions of a Unified Information Management. [...] it can be recognized that in recent years, BCEs have taken on a more specific, new, but still evolving meaning that refers to the storage and preservation of the organization's digital information or knowledge assets, either for compliance or as a source of revenue.'*

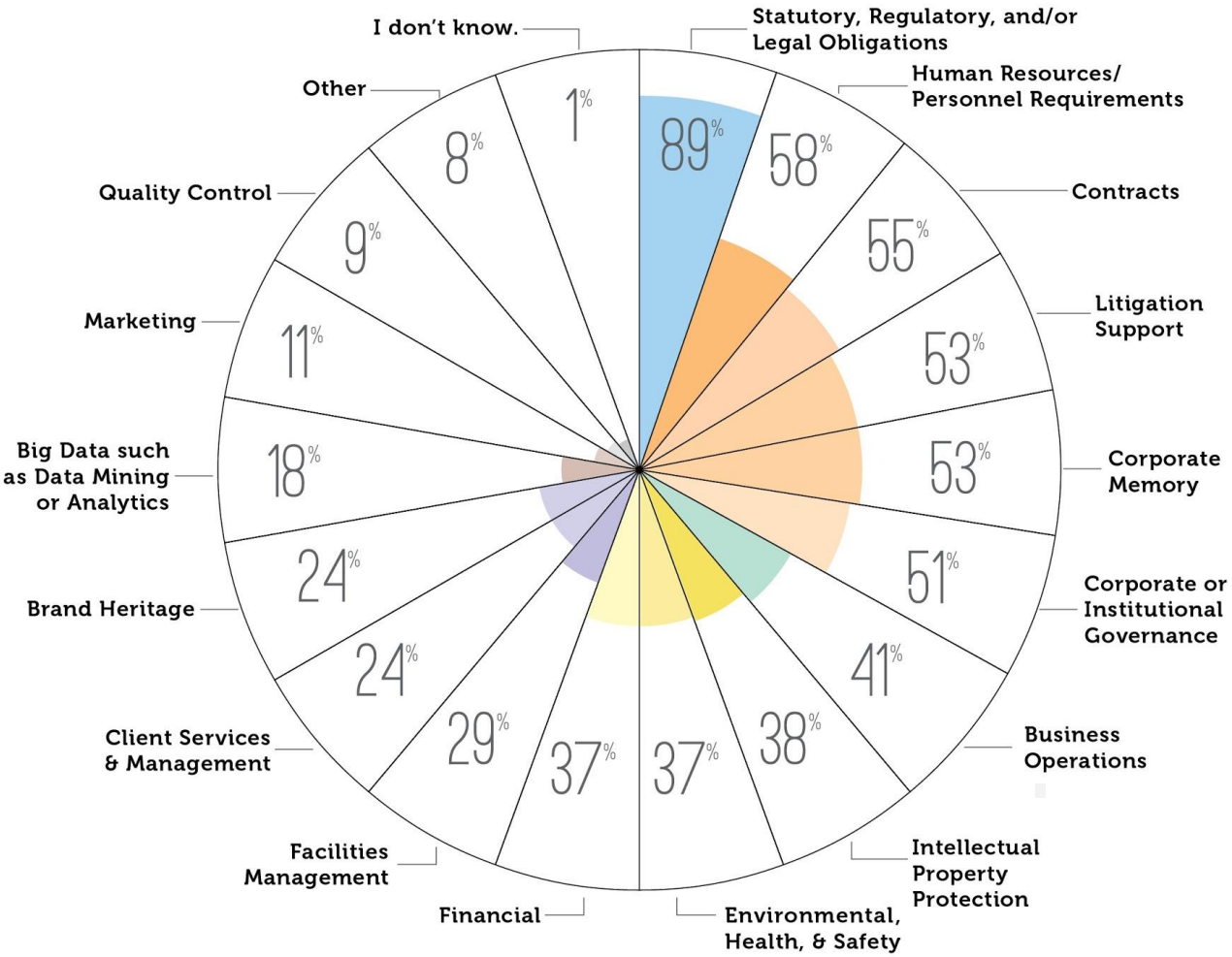
Source: Matthias L. Hemmje (2010) **Drivers for Digital Preservation**. (Online) Accessible: https://ercim-news.ercim.eu/en80/special/drivers-for-digital-preservation

Value propositions

Business Driver	Benefit
Collection development	Organize digital content as an institutional asset
Corporate memory	Maintain accumulated knowledge of the institution
User access	Ensure access to specific users over long term
Information re-use	Enable re-purposing and added-value use of digital data
Reputational protection	Safeguard institution's standing in the community
Legal & regulatory compliance	Comply with freedom of information, privacy, financial, health & safety
Business continuity	Eliminate data loss leading to catastrophic business disruption
Efficiencies & savings	Support a streamlined information management strategy
Protecting investment	Safeguard against commercial loss of digital content
Supporting digital ways of working	Future-proof and enable online-only ways of working with staff, customers and partners

Source: Adrian Brown (2013) **Practical Digital Preservation: a how-to guide for organizations of any size**. See: https://www.alastore.ala.org/content/practical-digital-preservation-how-guide-organizations-any-size

Why do we keep (digital) information?



Source: (detail from) Digital Preservation Coalition: **Executive Guide on Digital Preservation: Facts and Figures**. (Online), Available: https://dpconline.org/our-work/dpeg-home/dpeg-facts-and-figures

Policy and strategy frameworks

- *International*:
  UNESCO Recommendation concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form (2016); G20 Anti-Corruption Open Data Principles (2015); Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003); The FAIR Guiding Principles for Scientific Data Management and Stewardship (2016); [...]

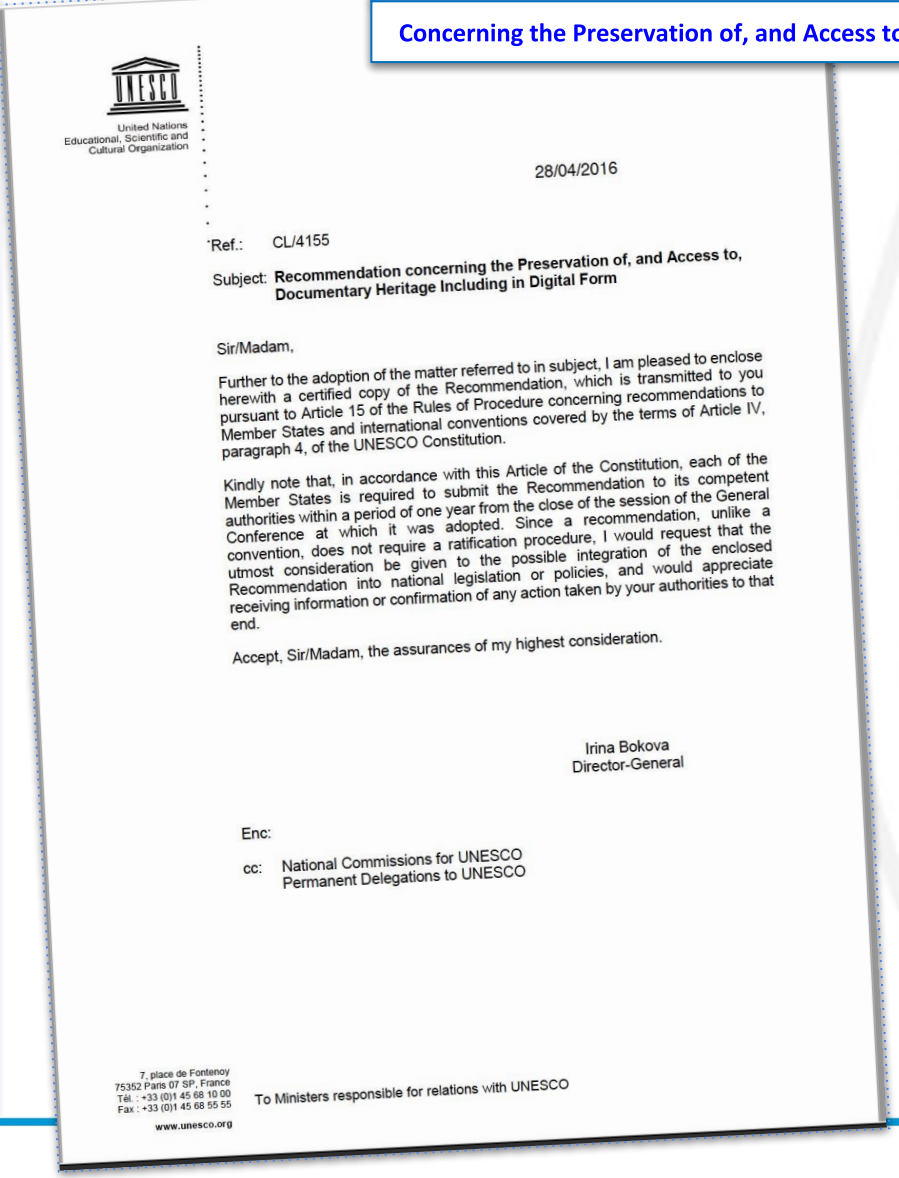
- *National*:
  National Integrated ICT Policy White Paper (2016); National Archives and Records Service Digitisation Strategy (2013); SAHRA / National Heritage Resources: Legislations and Regulations (1962 - 1999); Intellectual Property Rights from Publicly Financed Research and Development Act (2008); [...]

- *Institutional*:
  UCT Intellectual Property (IP) policy (2011); UCT Open Access (OA) policy (2014); UCT Research Data Management (RDM) policy (2017); [...]

- *Departmental*:
  UCT Libraries Digital Preservation Strategy (draft, 2019); UCT Libraries 'Horizon 2019' strategic plan (2014); [...]

UNESCO

Concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form (2016)



United Nations Educational, Scientific and Cultural Organization

28/04/2016

Ref.: CL/4155

Subject: **Recommendation concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form**

Sir/Madam,

Further to the adoption of the matter referred to in subject, I am pleased to enclose herewith a certified copy of the Recommendation, which is transmitted to you pursuant to Article 15 of the Rules of Procedure concerning recommendations to Member States and international conventions covered by the terms of Article IV, paragraph 4, of the UNESCO Constitution.

Kindly note that, in accordance with this Article of the Constitution, each of the Member States is required to submit the Recommendation to its competent authorities within a period of one year from the close of the session of the General Conference at which it was adopted. Since a recommendation, unlike a convention, does not require a ratification procedure, I would request that the utmost consideration be given to the possible integration of the enclosed Recommendation into national legislation or policies, and would appreciate receiving information or confirmation of any action taken by your authorities to that end.

Accept, Sir/Madam, the assurances of my highest consideration.

Irina Bokova
Director-General

Enc:

cc: National Commissions for UNESCO
Permanent Delegations to UNESCO

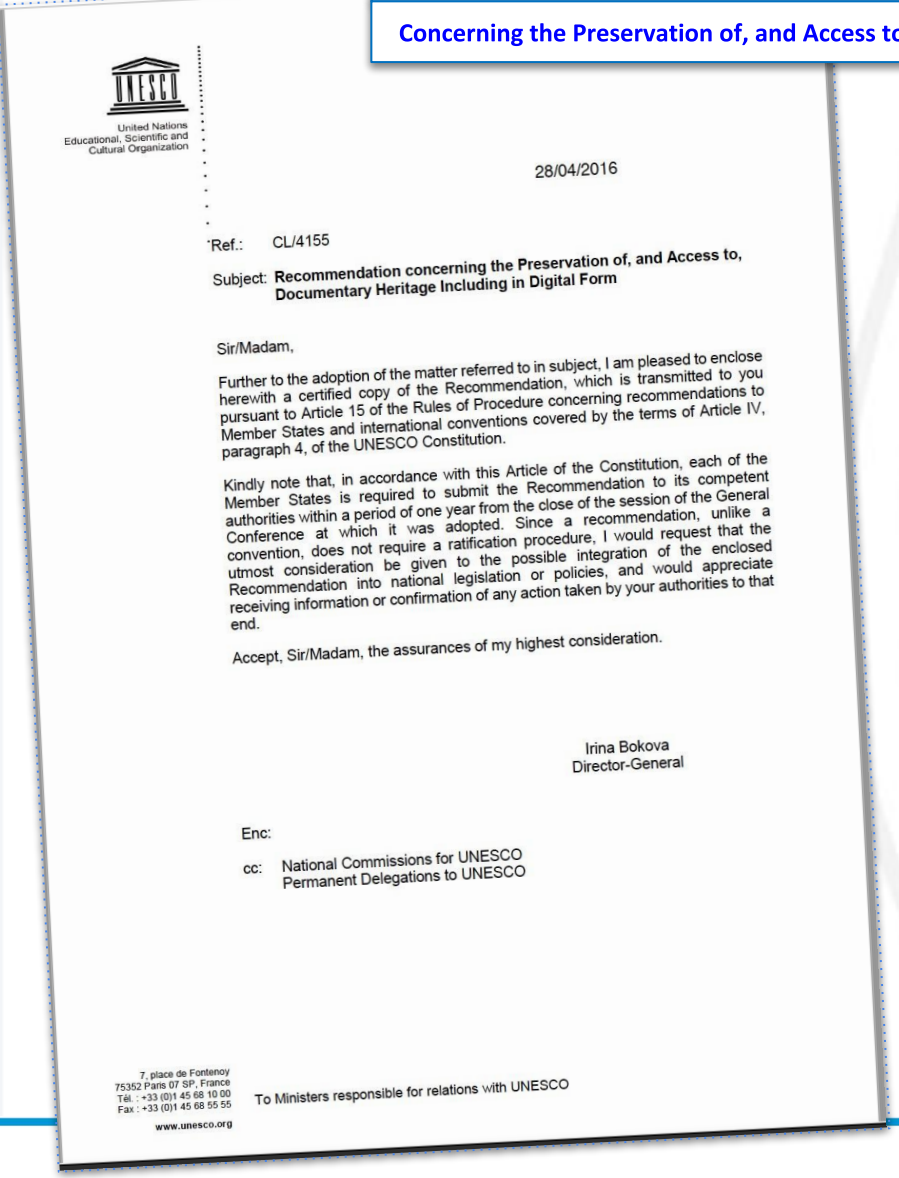
To Ministers responsible for relations with UNESCO

7, place de Fontenoy
75352 Paris 07 SP, France
Tél. : +33 (0)1 45 68 10 00
Fax : +33 (0)1 45 68 55 55
www.unesco.org

'[...] *documents produced and preserved over time, in all their analogue and digital forms through time and space, constitute the primary means of knowledge creation and expression, having an impact on all areas of humanity's civilization and its further progress, [...]* ***the preservation of, and long-term accessibility to documentary heritage underpins fundamental freedoms of opinion, expression and information as human rights*** [...]' (p.2)

UNESCO

Concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form (2016)



United Nations Educational, Scientific and Cultural Organization

28/04/2016

Ref.: CL/4155

Subject: **Recommendation concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form**

Sir/Madam,

Further to the adoption of the matter referred to in subject, I am pleased to enclose herewith a certified copy of the Recommendation, which is transmitted to you pursuant to Article 15 of the Rules of Procedure concerning recommendations to Member States and international conventions covered by the terms of Article IV, paragraph 4, of the UNESCO Constitution.

Kindly note that, in accordance with this Article of the Constitution, each of the Member States is required to submit the Recommendation to its competent authorities within a period of one year from the close of the session of the General Conference at which it was adopted. Since a recommendation, unlike a convention, does not require a ratification procedure, I would request that the utmost consideration be given to the possible integration of the enclosed Recommendation into national legislation or policies, and would appreciate receiving information or confirmation of any action taken by your authorities to that end.

Accept, Sir/Madam, the assurances of my highest consideration.

Irina Bokova
Director-General

Enc:

cc: National Commissions for UNESCO
Permanent Delegations to UNESCO

To Ministers responsible for relations with UNESCO

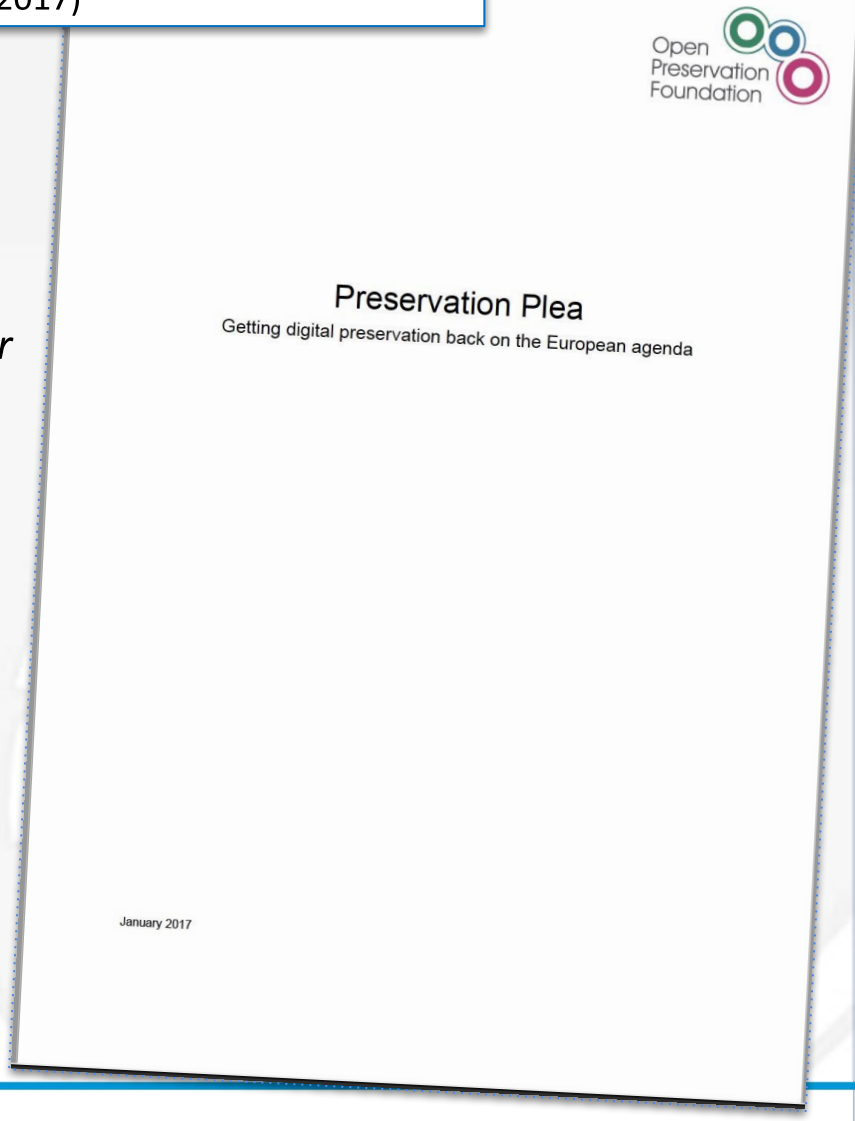
7, place de Fontenoy
75352 Paris 07 SP, France
Tél. : +33 (0)1 45 68 10 00
Fax : +33 (0)1 45 68 55 55
www.unesco.org

'The world's documentary heritage is of global importance and responsibility to all, and should be fully preserved and protected for all, with due respect to and recognition of cultural mores and practicalities. It should be permanently accessible and re-usable by all without hindrance. It provides the means for understanding social, political, collective as well as personal history. It can help to underpin good governance and sustainable development. For each State, its documentary heritage reflects its memory and identity, and thus contributes to determine its place in the global community.' (p.6)

Open Preservation Foundation

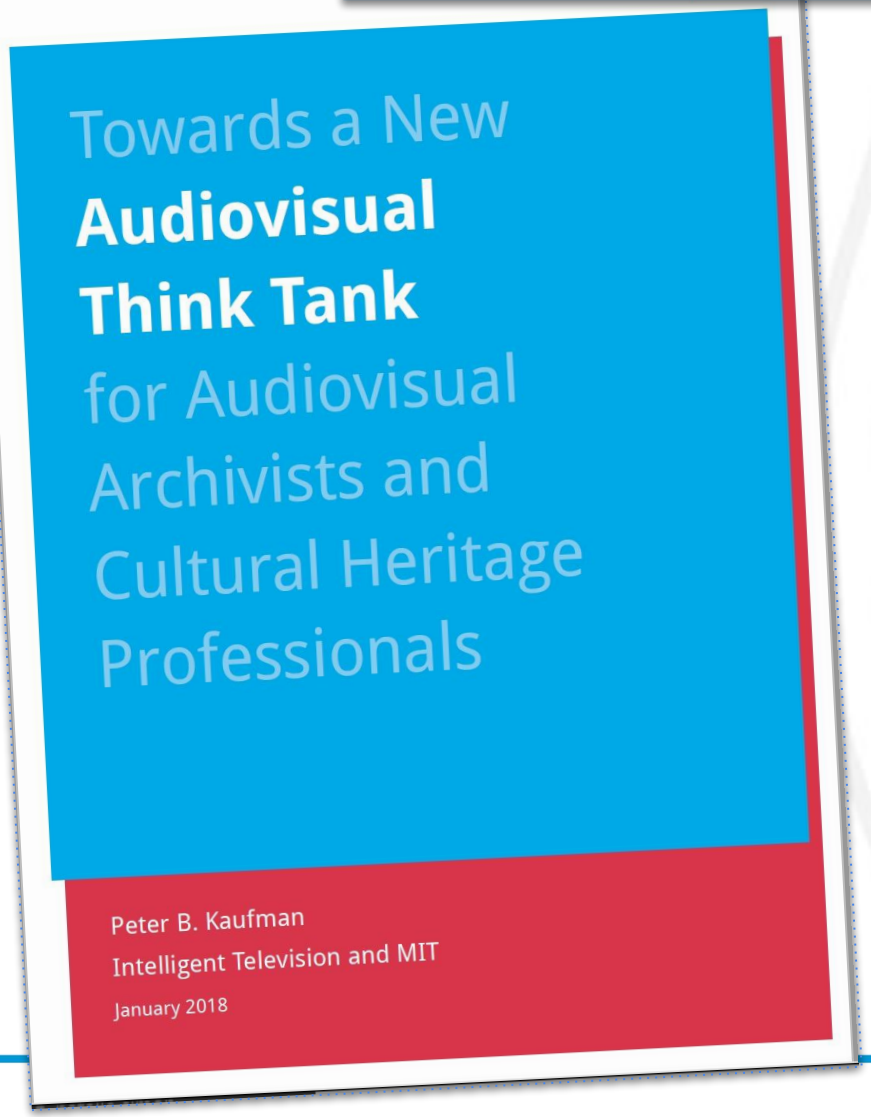
Preservation Plea (2017)

*'**Case of urgency**: In the near future the challenge in digital preservation is the chain of dependencies from the moment software is selected by our producers, digital information is generated, to when it is presented to our customers. In this chain it is our job to manage trustworthy and reusable information. In essence we adapt continuously to new technological environments by developing preservation tools and knowledge. This work is complex, time consuming and depends on a few individuals in science and heritage institutions. Consequently, we are concerned about the uptake, the cohesion with related topics and the sustainability of our preservation efforts, especially in light of the fact that preservation has not yet become a mainstream activity.'* (p.2)



Netherlands Institute for Sound and Vision

Towards a New Audiovisual Think Tank for Audiovisual Archivists and Cultural Heritage Professionals (2018)



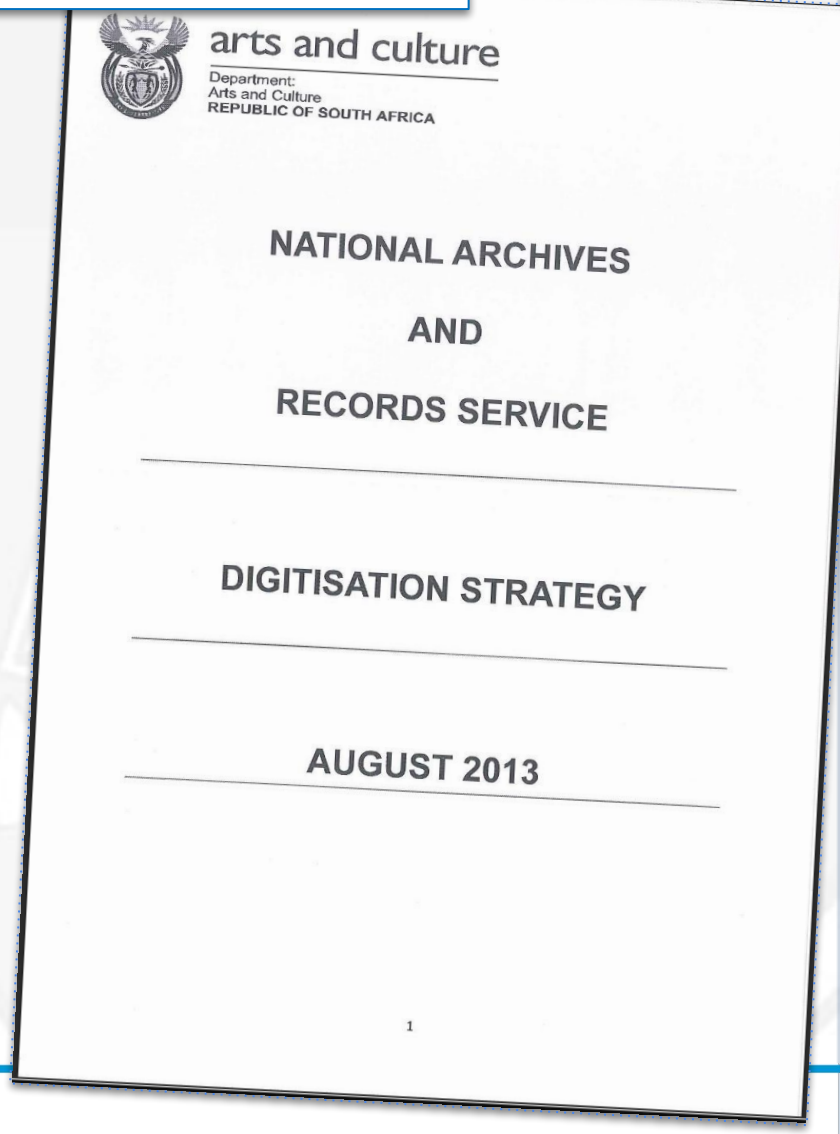
'Archiving thus is conducted not only on the back half, the defensive zone, of the field or pitch, as it were; it is also vitally proactive: it's productive publishing. It's not preserving alone, but also capturing and creating the human record – especially when activist groups are involved – and taking forward-facing steps to fulfill these urgent and giant social mandates and imperatives to remember, for today and for tomorrow.' (p.6)

S.A. National Archives and Records Service

Digitisation Strategy (2013)

'Preservation supports access and digitisation, and has the ability to further link the two. Provision of access is the driver for all our efforts to collect and preserve archival records.' (p.10)

'At international level the National Archives is a member of the International Council on Archives, who have developed the ICA/Atom which all members can use freely.' (p.12)


arts and culture
Department:
Arts and Culture
REPUBLIC OF SOUTH AFRICA

NATIONAL ARCHIVES

AND

RECORDS SERVICE

DIGITISATION STRATEGY

AUGUST 2013

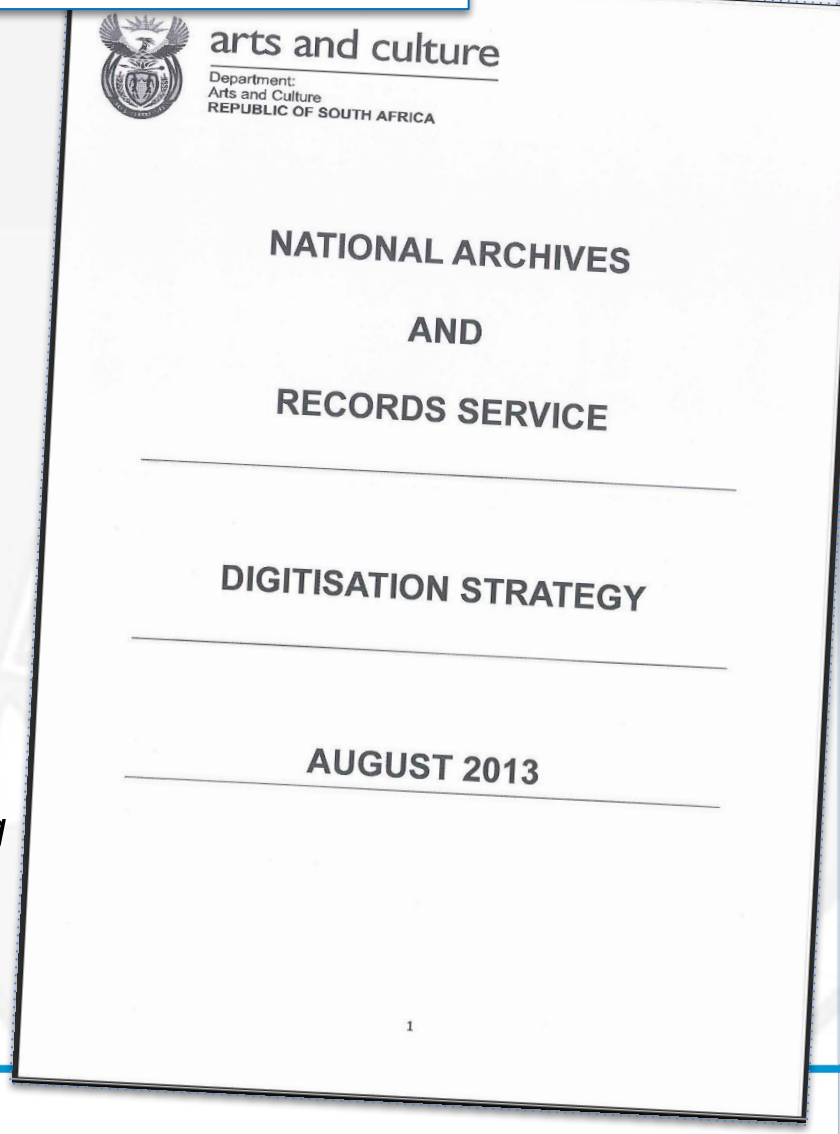
1

S.A. National Archives and Records Service

Digitisation Strategy (2013)

'The National Archives will ensure that the infrastructure and skills developed are sustainable in the long term. [...] The following skills must be developed for the sustainable management of digitisation programmes, projects and resources:

- *Advanced Computer Training*
- *Digital Resource Mgmt*
- *Project Mgmt*
- *Disaster Mgmt and Business Continuity*
- *Digital Capture and Photography/ scanning*
- *Systems Mgmt*
- *Digital Preservation'* (p.14)


arts and culture
Department:
Arts and Culture
REPUBLIC OF SOUTH AFRICA

NATIONAL ARCHIVES

AND

RECORDS SERVICE

DIGITISATION STRATEGY

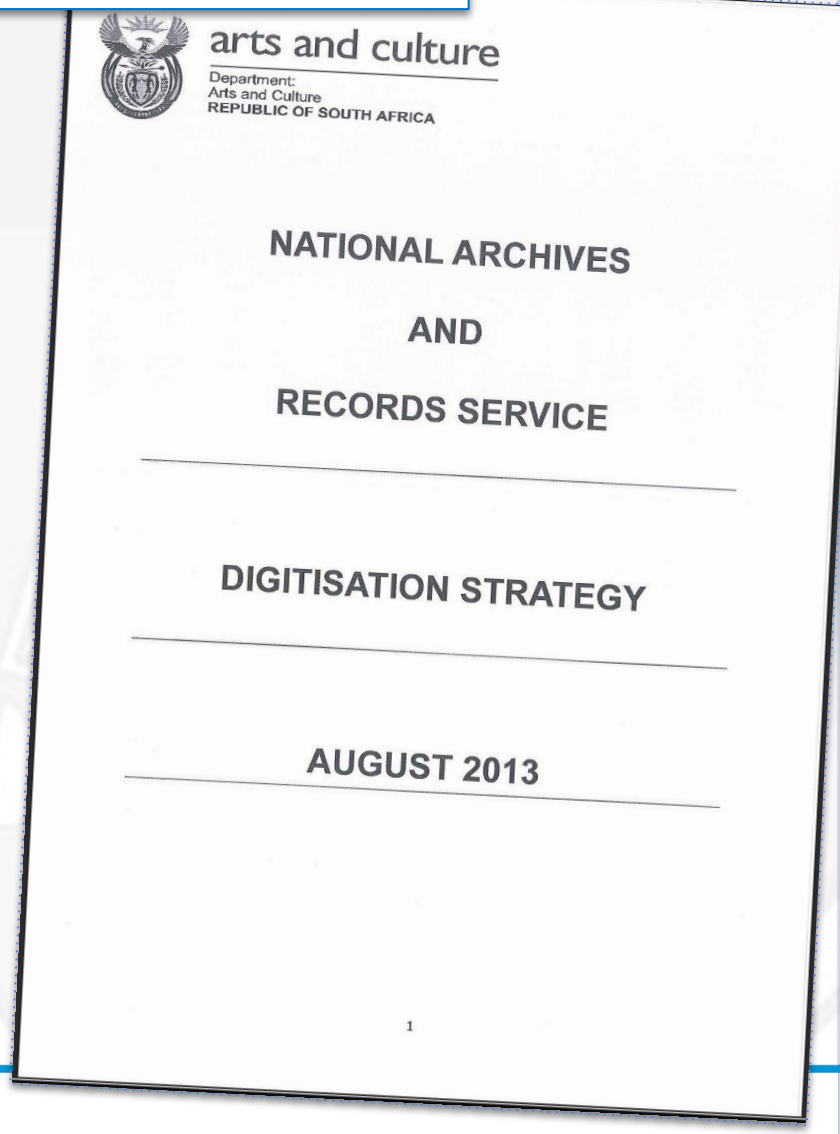
AUGUST 2013

1

S.A. National Archives and Records Service

Digitisation Strategy (2013)

‘ 10.7 ***Preservation Strategy****: The product of digitisation is a digital library or archives (resource). These resources have specific requirements for preservation which employs a proactive approach rather than a passive one. The biggest threats to digital resources are obsolescence, data integrity, trustworthiness, escalating storage costs and neglect. Once an organisation embarks on a digitisation programme, these threats become real and should be managed proactively. A dedicated section dealing with digital preservation will be considered for establishment. Digital preservation includes the management of born-digital material such as electronic records.* ***These are especially more vulnerable to loss in the absence of a proactive system to manage their preservation****. The National Archives has been identified as one of the National Digital Repositories in the National Policy on Digitisation of Heritage Resources.* [...] (p. 16)


arts and culture
Department:
Arts and Culture
REPUBLIC OF SOUTH AFRICA

NATIONAL ARCHIVES

AND

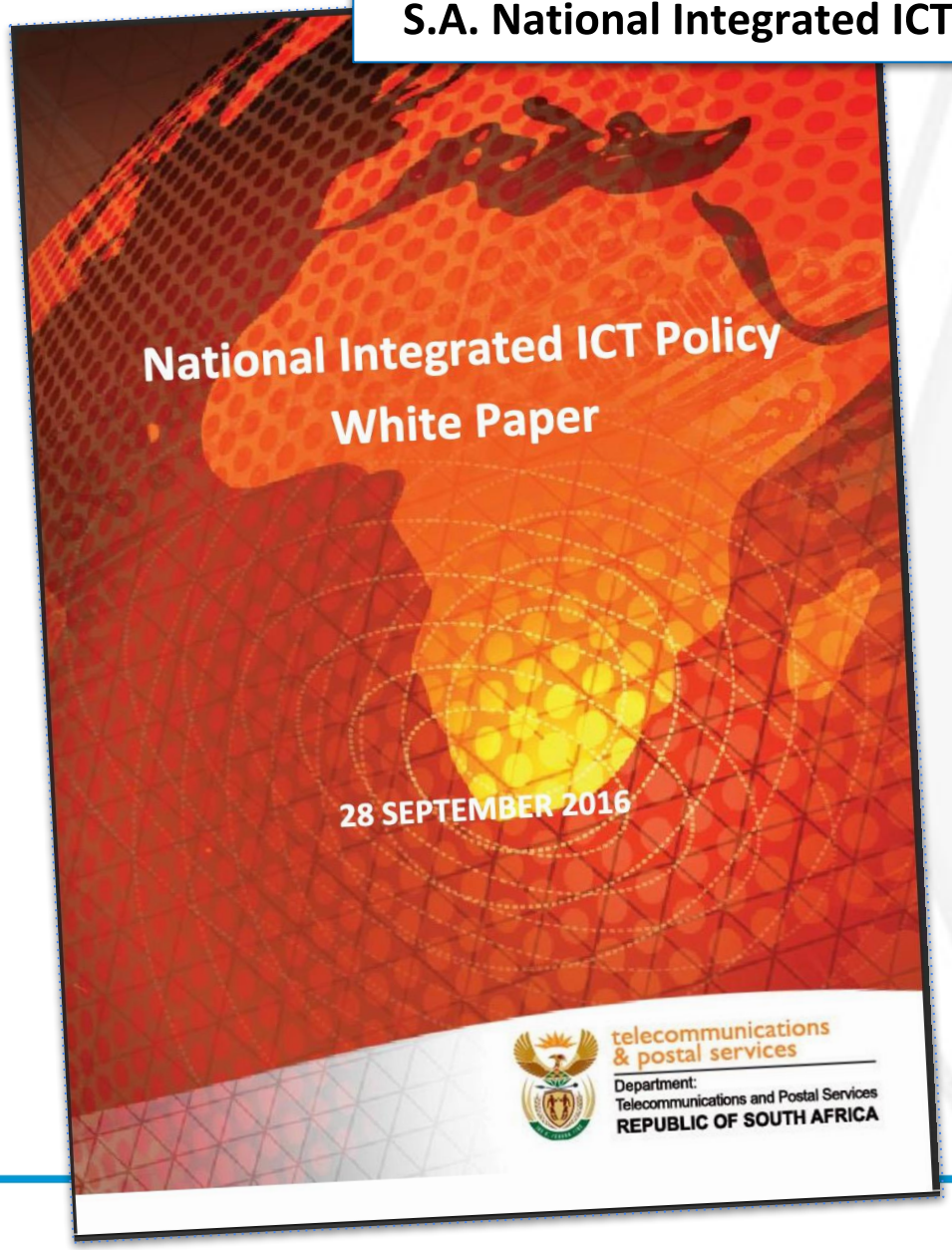
RECORDS SERVICE

DIGITISATION STRATEGY

AUGUST 2013

1

S.A. National Integrated ICT Policy White Paper (2016)

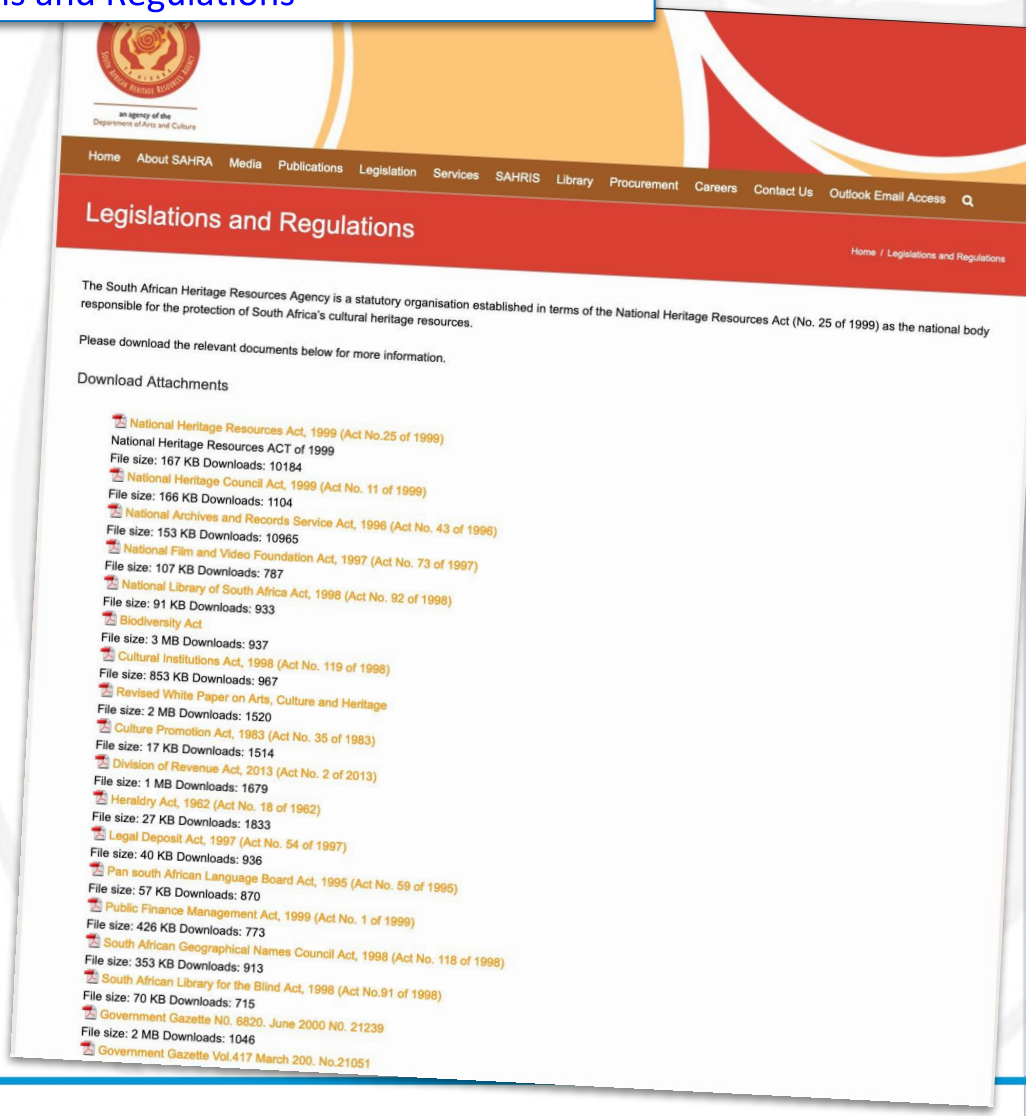


'***Digitising cultural heritage:*** *Government will also fast track existing programmes aimed at digitising cultural heritage (including content from libraries, museums, galleries, archives, film, music and video), and making this accessible online. Government will work together with other institutions, including private sector bodies and academic institutions to create a single access point for cultural heritage in South Africa.*' (p.133)

South African Heritage Resources Agency

Legislations and Regulations

South African Heritage Resources Agency (SAHRA) is a statutory organisation established under the National Heritage Resources Act, No 25 of 1999, as the national administrative body responsible for the protection of South Africa's cultural heritage. SAHRA, as an Agency of the Department of Arts and Culture (DAC), has been tasked to manage and implement the National Heritage Resources Act (NHRA) of 1999.



an agency of the Department of Arts and Culture

Home | About SAHRA | Media | Publications | Legislation | Services | SAHRIS | Library | Procurement | Careers | Contact Us | Outlook Email Access

Legislations and Regulations

Home / Legislations and Regulations

The South African Heritage Resources Agency is a statutory organisation established in terms of the National Heritage Resources Act (No. 25 of 1999) as the national body responsible for the protection of South Africa's cultural heritage resources.

Please download the relevant documents below for more information.

Download Attachments

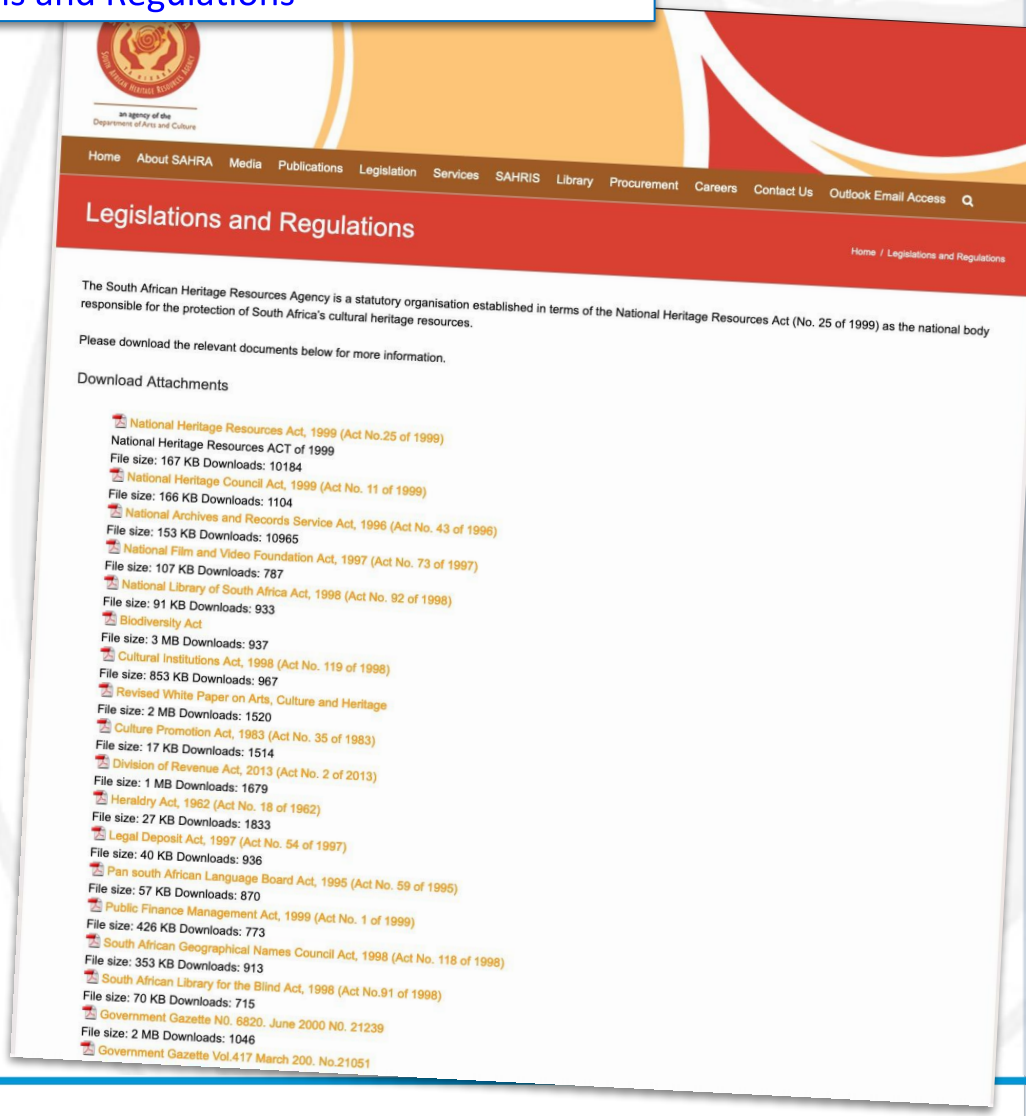
National Heritage Resources Act, 1999 (Act No.25 of 1999)
National Heritage Resources ACT of 1999
File size: 167 KB Downloads: 10184
National Heritage Council Act, 1999 (Act No. 11 of 1999)
File size: 166 KB Downloads: 1104
National Archives and Records Service Act, 1996 (Act No. 43 of 1996)
File size: 153 KB Downloads: 10965
National Film and Video Foundation Act, 1997 (Act No. 73 of 1997)
File size: 107 KB Downloads: 787
National Library of South Africa Act, 1998 (Act No. 92 of 1998)
File size: 91 KB Downloads: 933
Biodiversity Act
File size: 3 MB Downloads: 937
Cultural Institutions Act, 1998 (Act No. 119 of 1998)
File size: 853 KB Downloads: 967
Revised White Paper on Arts, Culture and Heritage
File size: 2 MB Downloads: 1520
Culture Promotion Act, 1983 (Act No. 35 of 1983)
File size: 17 KB Downloads: 1514
Division of Revenue Act, 2013 (Act No. 2 of 2013)
File size: 1 MB Downloads: 1679
Heraldry Act, 1962 (Act No. 18 of 1962)
File size: 27 KB Downloads: 1833
Legal Deposit Act, 1997 (Act No. 54 of 1997)
File size: 40 KB Downloads: 936
Pan south African Language Board Act, 1995 (Act No. 59 of 1995)
File size: 57 KB Downloads: 870
Public Finance Management Act, 1999 (Act No. 1 of 1999)
File size: 426 KB Downloads: 773
South African Geographical Names Council Act, 1998 (Act No. 118 of 1998)
File size: 353 KB Downloads: 913
South African Library for the Blind Act, 1998 (Act No.91 of 1998)
File size: 70 KB Downloads: 715
Government Gazette N0. 6820. June 2000 N0. 21239
File size: 2 MB Downloads: 1046
Government Gazette Vol.417 March 200. No.21051

South African Heritage Resources Agency

Legislations and Regulations

SAHRA aim to fulfil their mandate of promoting social cohesion in S.A. by:

- Identification, conservation and management of heritage resources in South Africa so that they can contribute to socio-economic development and nation building;
- Developing of norms, standards and charters for the management of heritage resources in South Africa and codes of international best practice; and
- Contributing to skills and knowledge production and transformation in heritage resources management in South Africa and beyond.



World Digital Preservation Day
7 November 2019

DIGITAL LIBRARY SERVICES

Practicalities of Digital Preservation

- *Best Practices*
- *Tools & Systems*
- *Vendors*
- *Roles*
- *Activities*

Digital Preservation **best practices**

U.S. National Archives (NARA) **Digital Preservation Framework** on Github

news!

The framework comprises an instrument for assessing file format risk, the results of NARA's assessment of over 350 file format variants, and draft preservation plans for those formats.

We are inviting comments and discussion on the plans using the Github Issues feature:

- *What revisions can you suggest to the proposed processing and preservation actions for the formats?*
- *Are the Essential Characteristics for each record type comprehensive enough for digital preservation?*
- *Are the proposed preservation actions for the formats technically appropriate?*
- *Are there appropriate tools for processing and preservation of specific formats that we do not have listed?*
- *Are there other high priority formats that we need plans for?*
- *What can you suggest in terms of appropriate public access versions of the formats?*

Source: https://github.com/usnationalarchives/digital-preservation

Digital Preservation **tools & systems**

UK National Archives > Digital preservation tools and systems

Your approach to Digital Preservation should be modular and flexible, to ensure it is sustainable. A combination of tools and technology that are currently available may be the most cost-effective means of achieving this flexibility:

- *file format identification from The National Archives: DROID (freeware)*
- *file registry database from The National Archives: PRONOM (freeware)*
- *Community Owned Digital Preservation tool registry: COPTR*
- *SPRUCE project, including a Digital Preservation Business Case toolkit: SPRUCE*

There are a growing number of options available on the market that address the digital preservation needs of archives. We have compiled a list of digital preservation solutions for archives, including commercial and open source. Product descriptions have been taken in part or in whole from the product web pages. For each option we have included information such as:

- *the features of the product*
- *whether it is OAIS compliant*
- *whether it is open source or commercial*
- *whether it is software or a service*
- *whether there is easy to access documentation and an active user community (NB for open source products)*
- *whether storage is included*
- *what level (if any) of file encryption is offered*

Source: https://www.nationalarchives.gov.uk/archives-sector/advice-and-guidance/managing-your-collection/preserving-digital-collections/digital-preservation-tools-systems/

Digital Preservation **vendors**

An **extensive evaluation process** of digital preservation platforms was conducted by DLS for UCT Libraries, which included (amongst others), **Archivematica & AtoM** (Open Source as well as hosted versions) (Artefactual Systems), **Rosetta** (ExLibris), **Preservica** (Arkivum), and **Preservica**.

archivematica®

atom

Preservica
Active digital preservation

arkivum
Bringing archived data to life

ExLibris Rosetta

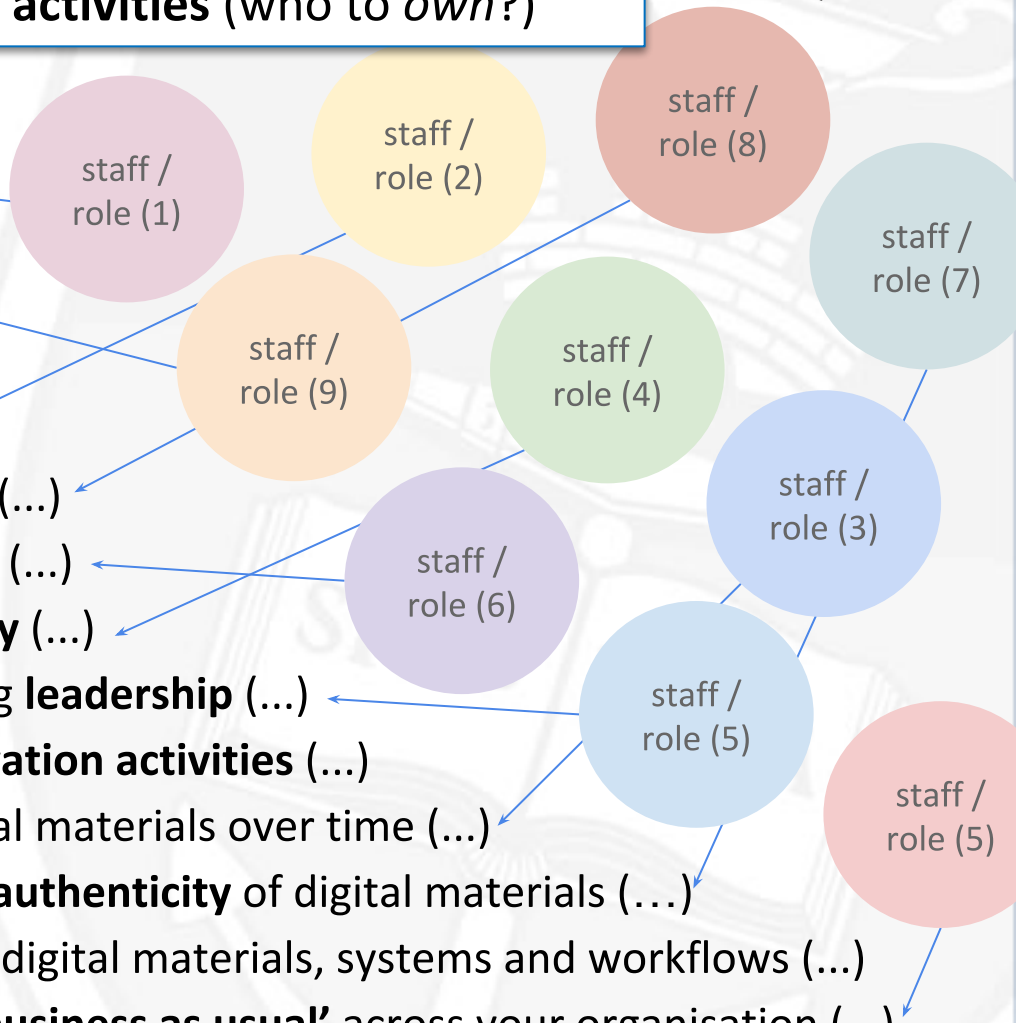
The evaluation compared **open source** and **licensed options**, and took into close consideration both **data sovereignty** (i.e. requirements for on-premise storage) and **local infrastructure support staffing costs**.

Digital Preservation **roles**

- **Budgets** are **limited** (austerity measures), and there is **not enough IT staff** at most SA HEIs (Faculties, Libraries, IT services) **to host independent, local solutions**, **including so-called 'free' Open Source** platforms.
- **Open Source requires ongoing local, specialist IT support**, including: needs analysis, liaison, implementation, customisation, development, upgrades, migration, etc.
- **Federated licensing** (TENET?) would be preferable to **individual institutional licenses**.
- Organisational structures require urgent review to embed new**, specialised skills**:
  - **Data Analysis, Mining,** and **Visualisation** (Digital Humanities; GIS; ...)
  - **Data Archiving** (File format conversion; Bit curation; DP; …)
  - **Data Curation** (Data access, archiving and publishing; Metadata schemas; …)
  - **Data Librarianship** (Data acquisition, citation, location and re-use; ...)
  - **Digital Scholarship** (Liaison; Open Scholarship; R&D; Digital Humanities; …)
  - **Repository management** (Coding; Digital library infrastructure; Semantic web; ...)
  - **Research Data Management** (Funder mandates; Advocacy; Support services, …)

Digital Preservation **activities** (who to *own*?)

- Capture **metadata** (...)
- **Liaise** with **stakeholders** (...)
- Use appropriate **standards** (...)
- Provide appropriate **access** (...)
- Carefully **appraise** and **select** (...)
- Provide supporting **documentation** (...)
- Keep up with technological **changes** (...)
- Plan and develop **strategy** and **policy** (...)
- **Work together** with strong, enabling **leadership** (...)
- Assign appropriate **levels of preservation activities** (...)
- **Add value** to an organization's digital materials over time (...)
- Ensure the continued **integrity** and **authenticity** of digital materials (…)
- Actively monitor, plan, and manage digital materials, systems and workflows (...)
- Help make digital preservation be **'business as usual'** across your organisation (...)



Adapted from: Digital Preservation Coalition: **Executive Guide on Digital Preservation for all organizations: All organisations**. (Online), Available: https://dpconline.org/our-work/dpeg-home/dpeg-organisation-type/dpeg-all-orgs

World Digital Preservation Day
7 November 2019

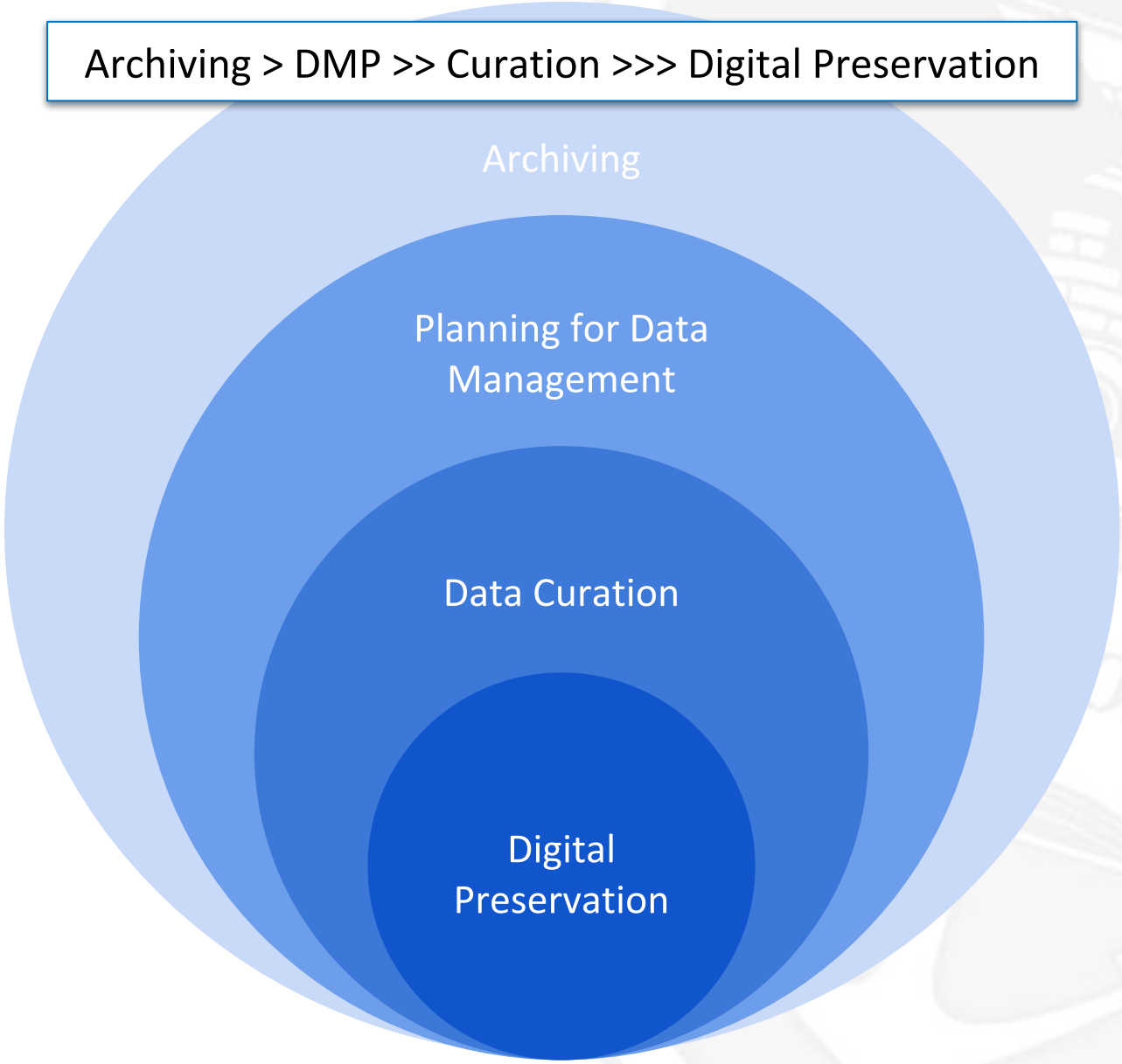
DIGITAL LIBRARY SERVICES

A Data Lifecycle approach to Digital Preservation

- **Plan & Design**: UCT DMP
- **Collect & Capture**: Digitisation for preservation & access
- **Collaborate & Analyse**: UCT Open Science Framework
- **Discover, Reuse & Cite**: selected repositories (see: re3data)
- **Share & Publish**: *to-be-named-still* (Omeka-S)
- **Manage, Store, Preserve**: *to-be-named-still* (Arkivum Perpetua)

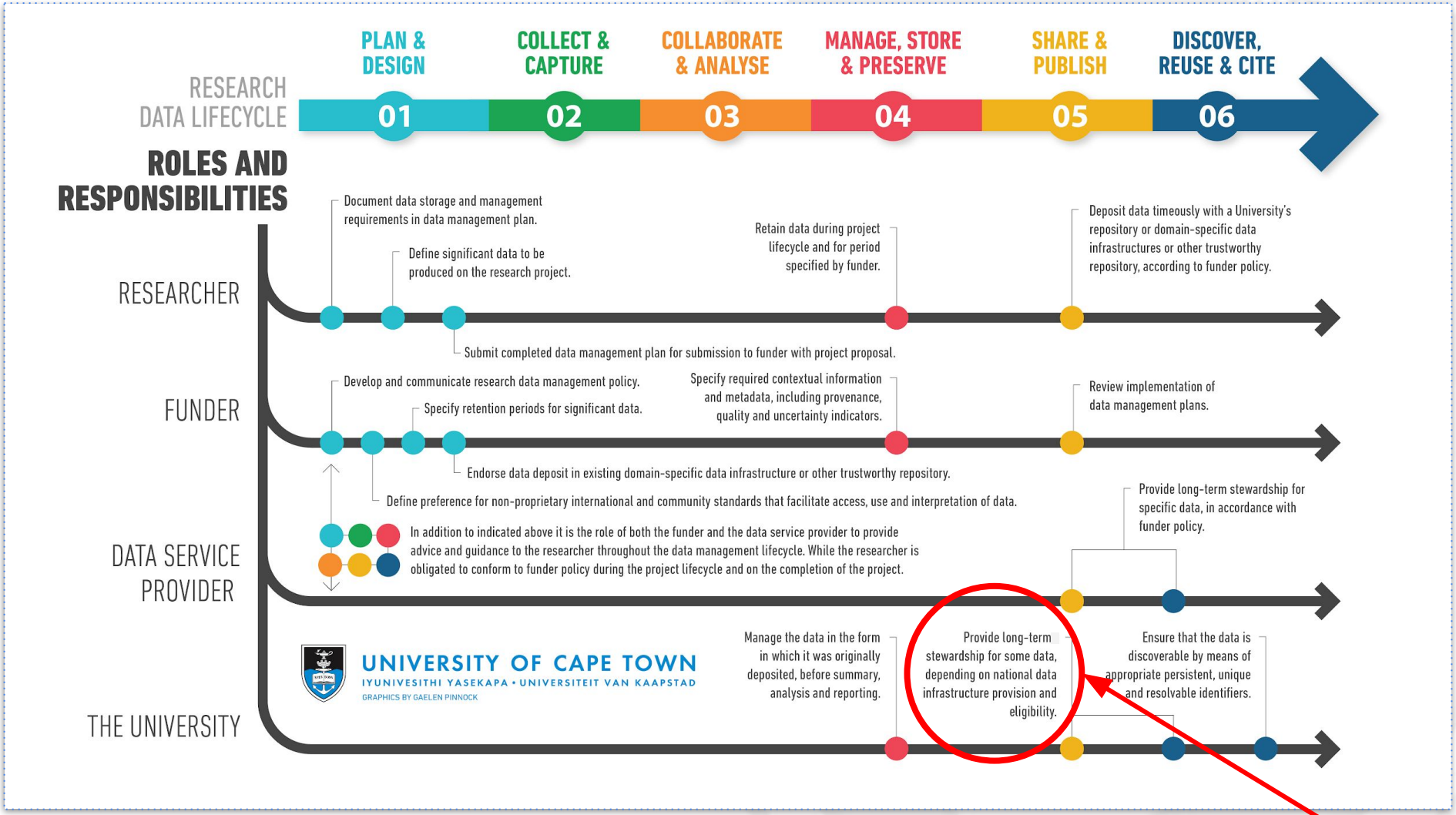
The (R)DM Lifecycle

Archiving > DMP >> Curation >>> Digital Preservation



DIGITAL LIBRARY SERVICES

RDM **roles & responsibilities** at UCT



Source: UCT RDM **Roles and Responsibilities**: https://upload.wikimedia.org/wikipedia/commons/0/04/UCT_RDM_Roles-and-Responsibilities.png

PLAN & DESIGN

Plan & Design

Planning for your various needs upfront, you are able to resource them appropriately. Collaborate with all your stakeholders at the budget application stages in order to strategically approach all the phases of digital preservation.

- ★ design digitisation & preservation projects
- ★ plan consent for publishing (IP / copyright, release forms, etc.)
- ★ locate & (re-)process existing data
- ★ plan data management (formats, storage, etc.)
- ★ digitisation to intl. standards and best practice
- ★ capturing of technical metadata
- ★ creation of descriptive metadata

Typical DMP **questions**

- **What type of data** will be generated?
- How will the data be **named and referenced**?
- What **file formats** are involved?
- What data and **metadata standards** will be followed?
- Who will **have access** to the data?
- How and when will the data be **shared**?
- Will the data be **digitally preserved**? If yes, how so?
- How will the data be **licensed**?
- How will **privacy** or **confidentiality** be ensured?

Adapted from: OSF Guides > Best Practices > Handling Data > **Creating a data management plan (DMP)**. Available: http://help.osf.io/m/bestpractices/l/618674-creating-a-data-management-plan-dmp

Planning for **Data Management**

UCT DMP

DIGITAL LIBRARY SERVICES

Plan | Process | Enhance | **Preserve** | **Publish** | Reuse

Proposal

MoU

DMP
(*proposal*)

"We will do XYZ to manage our data"

resourcing

DMP
(*active*)

"We are doing XYZ to manage our data"

documentation / review

DMP
(*final*)

"We have done XYZ to manage our data"

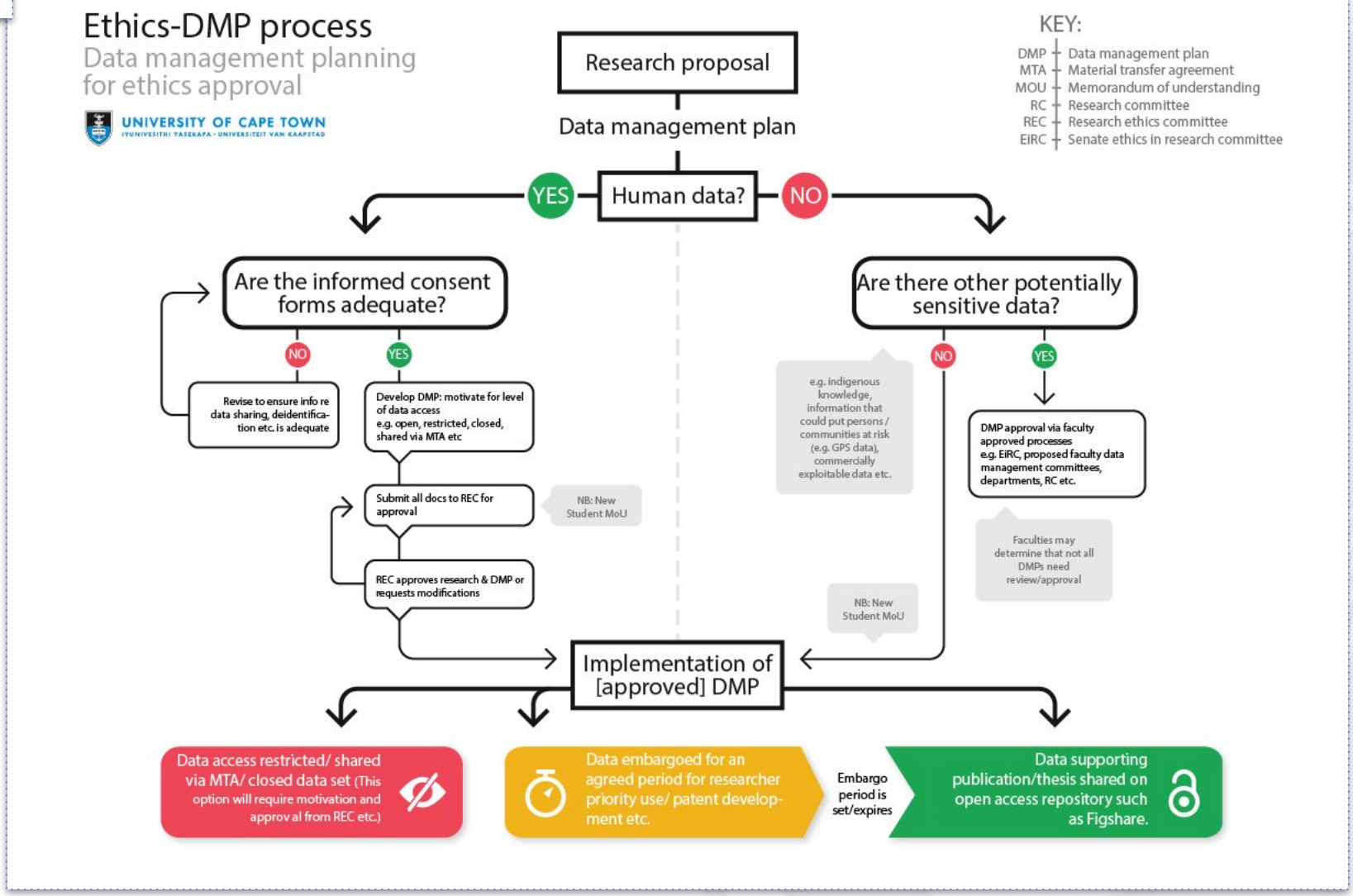
UCT DMP

arkivum
Bringing archived data to life

iiif

omeka s

Data Management Planning & **Ethics** at UCT



Source: https://commons.wikimedia.org/wiki/File:UCT_RDM_DMP-for-ethics-approval.png

COLLECT & CAPTURE

Collect & Capture

Obtaining and/or creating data often requires the use of complex tools and software, which collect raw data that then need to be stored and managed. Ensuring that data are captured and stored effectively involves setting up relevant soft- and hardware, as well as NB: well-documented, cross-departmental workflows.

- ★ creating / logging data
- ★ checking, validating, and cleaning data
- ★ managing active data storage
- ★ enhancing metadata
- ★ transcription, translation
- ★ anonymise / de-identify data (where necessary)

‘Collect & Capture’ actions

Special Collections as data producers

While collecting and capturing (raw) data, also capture and created accurate, meaningful metadata:

- Documentation of **data types** generated:
  - Form: *What kind of data does it hold?*
  - Stability: *How does it change over time?*
- Documentation of **identifiers** used:
  - File naming conventions, e.g. barcodes
  - Data dictionary
- Documentation of (DP) **process**:
  - *readme*.txt to be deposited with data (AIP)

Adapted from: OSF Guides > Best Practices > Handling Data > **Creating a data management plan (DMP)**. Available: http://help.osf.io/m/bestpractices/l/618674-creating-a-data-management-plan-dmp

Digitisation for Digital Preservation & Access

e.g.: http://www.digitalservices.lib.uct.ac.za/dls/what-we-digitise

Some requirements, by example ...

Example: digital preservation of video tapes:

Requirements (overall): *Specialised equipment* (legacy and current) + *Scarce skills*

1) **Legacy media** (analogue or digital)
   e.g.: film collections (NB: 20+ formats at SC!).
2) **Documentation** regarding mandates, rights, description, QA, delivery
   e.g.: Digital Preservation Policy; AV Digitisation guideline; service manuals.
3) **Physical storage**, cleaning and preparation facilities, equipment, materials and skills
   e.g.: cold room, tape cleaning/conditioning machines (per format), training.
4) **Legacy playback equipment**; knowledge and skills to operate & troubleshoot
   e.g.: playback machines (per format), training, servicing/repair.
5) Up to date **digitisation equipment** (DACs), capture software (DAW) and skills
   e.g.: video capture cards, fast computer and HDD, video software, training .
6) Large, fast **repository solutions** for storage, management, preservation and access
   e.g.: Media Asset Management system, high-availability server; systems and processes for preservation (size!) and online exhibition (NB: streaming!).

COLLABORATE & ANALYSE

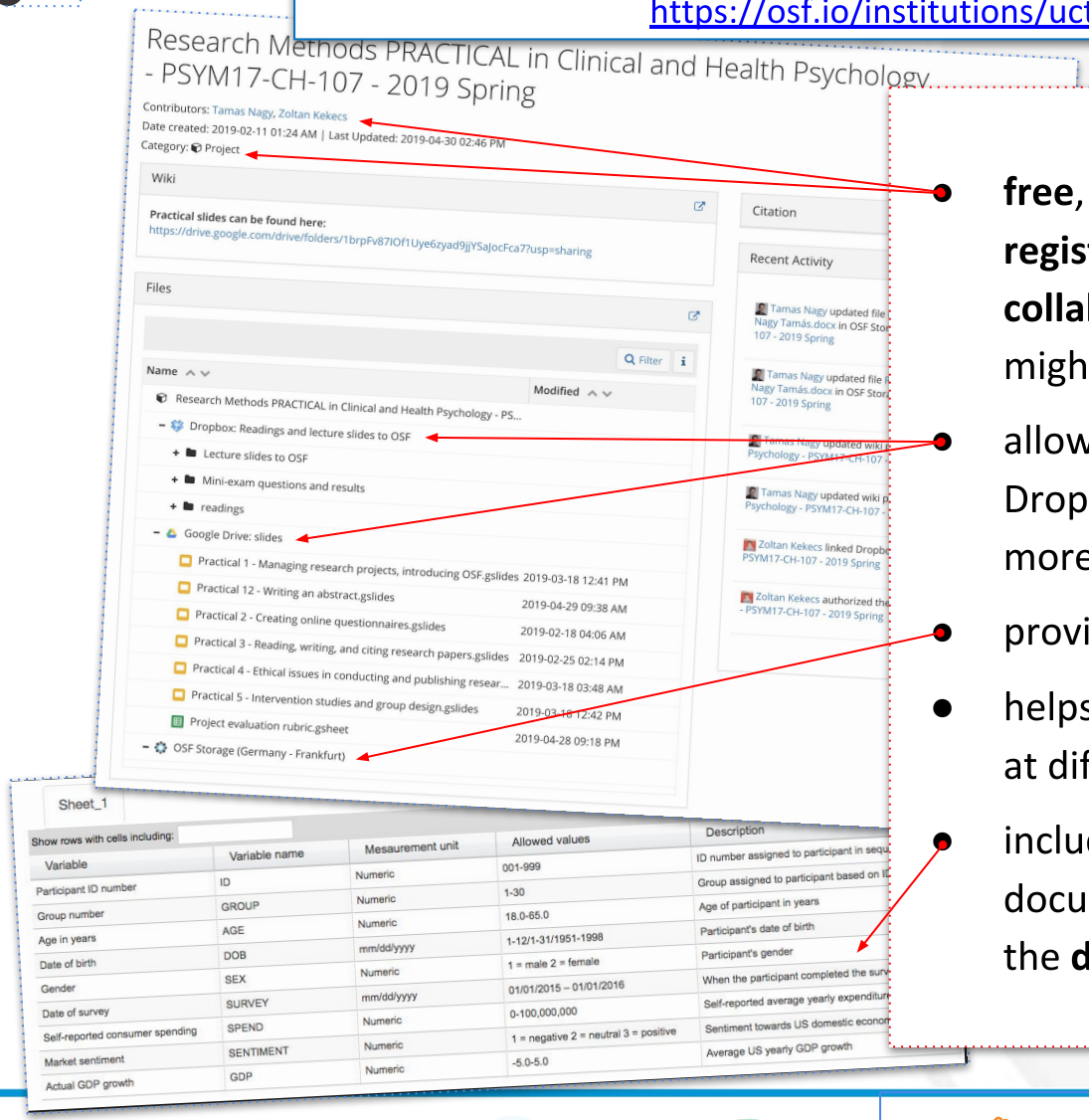
Collaborate & *Process*

Archival projects today often include teams scattered geographically, who need access to the same data at the same time. Consider creating shared data stores, dataset transfers, file sharing and other facilities or software required for effective collaboration, including data processing (e.g. virtual machines, cloud-based software etc.).

- ★ checking, validating, and cleaning data
- ★ creating / enhancing metadata; transcription; translation
- ★ anonymising / de-identifying data
- ★ describing, managing and storing data
- ★ prepare data products for preservation and publication

Open Science Framework (OSF)

https://osf.io/institutions/uct/



- **free**, online platform that allows you to **register** your project, **manage collaborators,** and **centralise data** that might be stored at different locations
- allows **integrations** with Google Drive, Dropbox, OneDrive, figshare, and many more
- provides **unlimited, free storage**
- helps with **creating versions** of your project at different stages ('forking')
- includes **wiki-components** for ease of documentation and description, including the **development of a data dictionary**

OSF integrations along the data life-cycle

Advanced digital scholarship

Data Analysis and Mining:

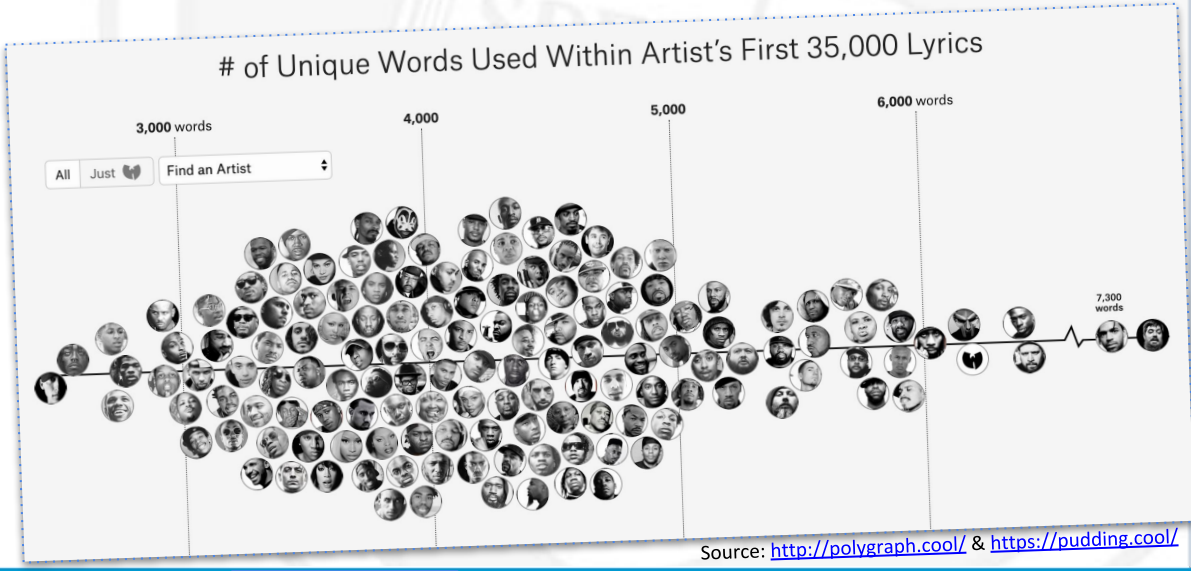
Tools that help you identify patterns in large volumes of data, combining statistics, AI and machine learning.

- Tools and processes for data de-identification, to safeguard privacy of patients.
- Tools and process for text analysis.

Source: **Tableau**: https://www.tableau.com/

Data Visualization:

Tools that develop a graphical presentation of data and information through visual means.



Source: http://polygraph.cool/ & https://pudding.cool/

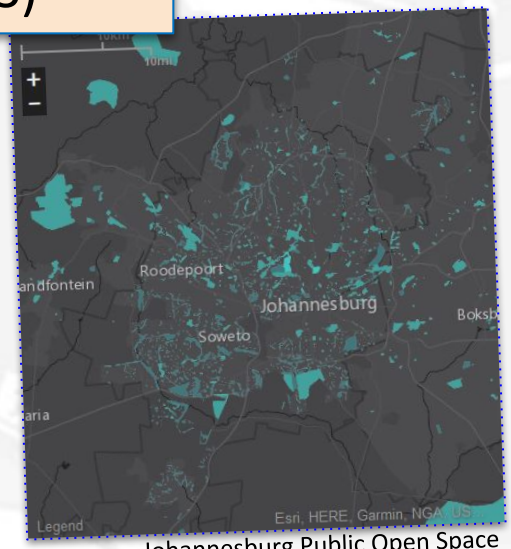
Digital Humanities:

Tools, processes and critical awareness found in the intersection between digital technologies and fields of study within the humanities.

Geographic Information Systems (GIS)

Everything Happens Somewhere:

- Because everything happens somewhere everything can be associated with a spatial location.
- These locations can be mapped in space, either for simple visualisation or for complex analyses.

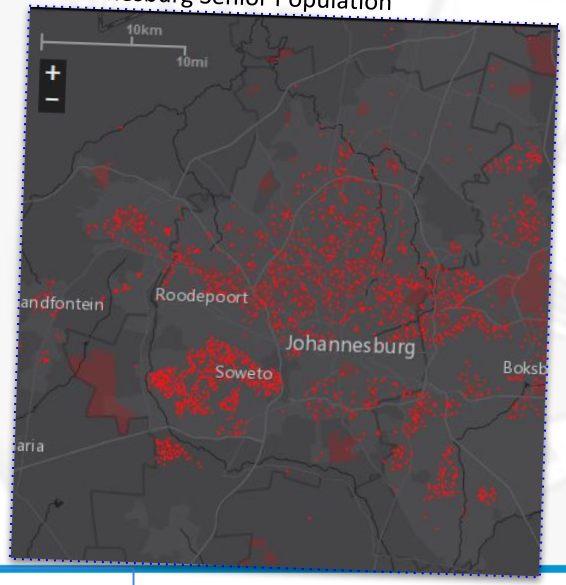


Johannesburg Public Open Space

Data Visualisation (Maps):

- Maps are an incredibly powerful visualisation tool which allow us to view and display our data in interesting and informative ways. They allow us to see patterns in our data, not just find them.
- They also allow us to communicate our findings in a clear and succinct manner.

Johannesburg Senior Population



Images sourced from UrbanObservatory.org's App

Geographic Information Systems (GIS)

Data Analysis (Making Information):

The full potential of GIS is realised when performing spatial analyses. Different types of analyses exist to satisfy various needs:

- **Overlay** Analysis allows us to compare different data types, e.g. Mean Annual Rainfall and Crop Type.
- **Geostatistical** Analysis allows us to perform statistical analyses of correlated spatial data, e.g. Hotspot Analysis.
- **Network** Analysis allows us to calculate travel times and service delivery areas, e.g. "Golden Hour" coverage or Clinic's Service Area.
- **Dashboards** of real time sensor feeds for live monitoring, e.g. Resource Usage; Traffic Volumes; Fleet Management.

Difference in Malaria rates between 2000 and 2015.
From the urbanobservatory.org

DLS' GIS services assist with GIS software acquisition, project planning, troubleshooting, analysis and cartographic design. More @ www.gis.uct.ac.za

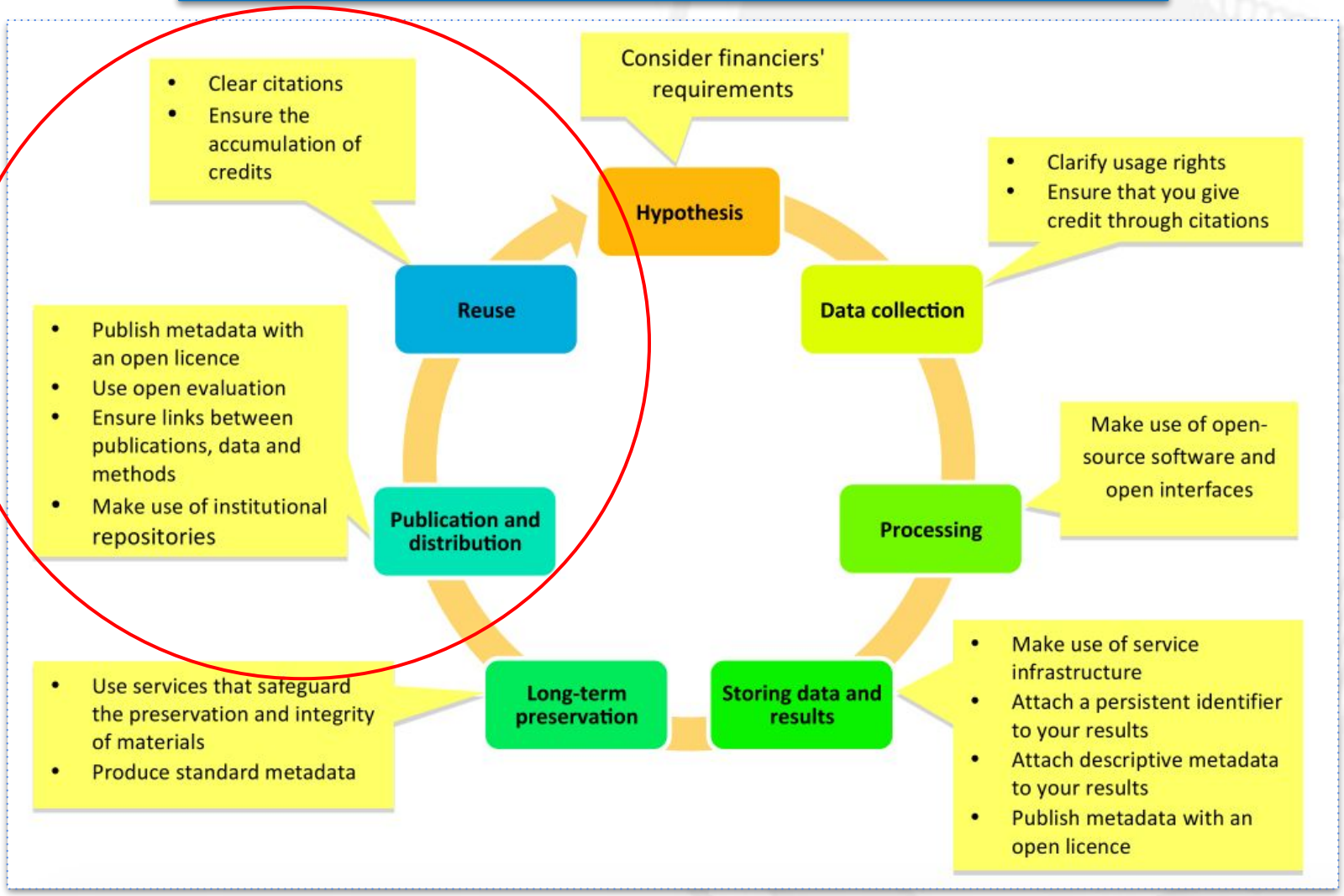
DISCOVER, REUSE & CITE

Discover, Reuse & Cite

The outputs of digitisation / digital preservation projects should be publicly available as citable, scholarly resources. Open-access publishing of data should go hand-in-hand with other OA publishing efforts.

- ★ follow-up research
- ★ new research
- ★ research interviews
- ★ crowdsourced annotations / public feedback
- ★ teaching and learning

Open discovery, reuse and citation



Source: Foster Open Science: **What is Open Science?** Figure 1. **Promoting openness at different stages of the research process.** https://www.fosteropenscience.eu/content/what-open-science-introduction

SHARE & PUBLISH

Share & Publish

Scholarly resources should be '*as open as possible, as closed as necessary*,' i.e. publicly available on an open-access platform.

★ manage copyright
★ distribute data
★ control access
★ promote data

- Guidance on publishing and sharing sensitive data:
  - http://www.ands.org.au/guides/sensitivedata
- Tools for anonymisation, de-identification, and disclosure control:
  - Amnesia (OpenAIRE)
  - Statistical Disclosure Control (sdcMicro) (IHSN)

Consider: 5 ★ **Open Data** {Tim Berners-Lee}

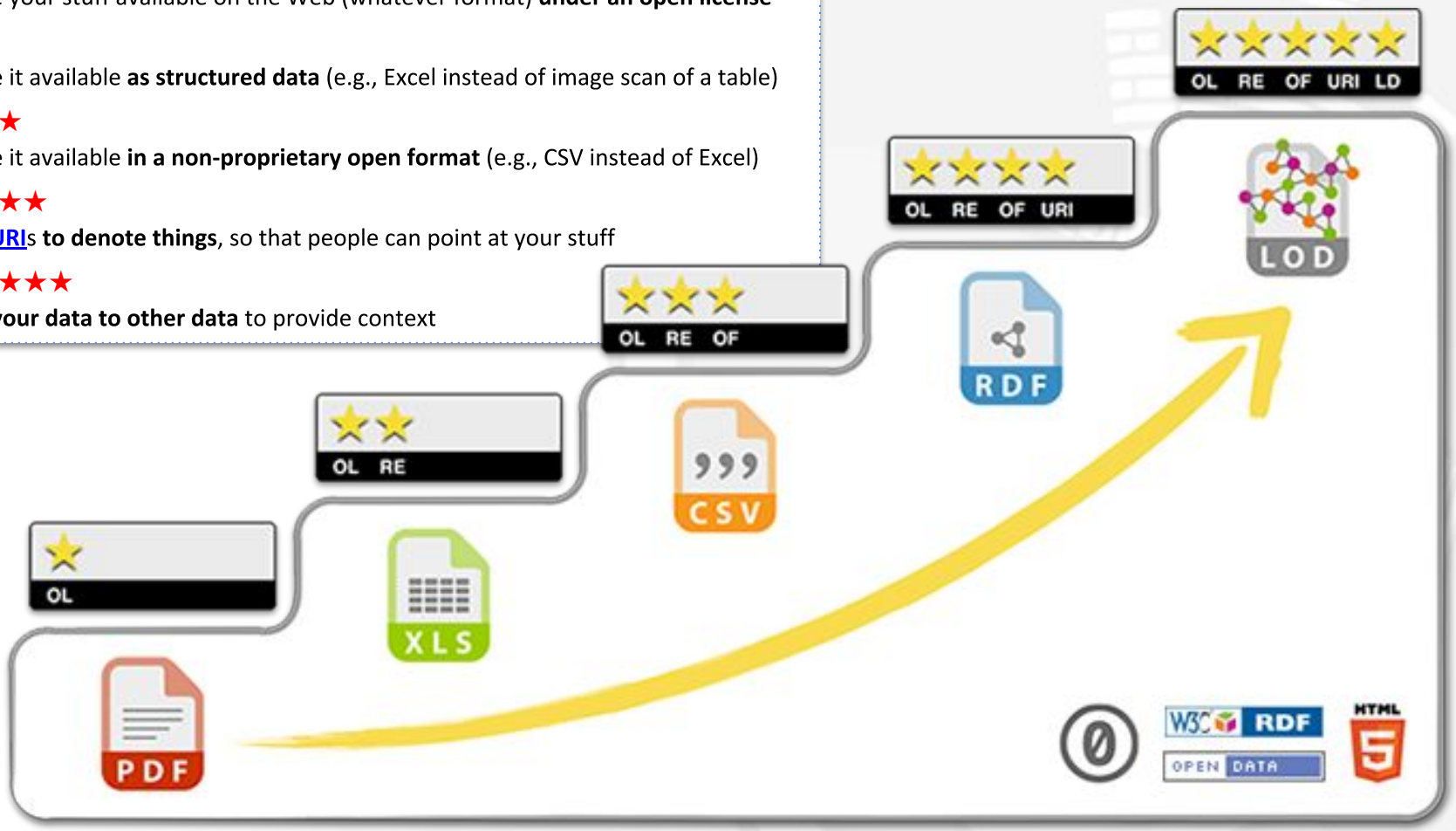
★
make your stuff available on the Web (whatever format) **under an open license**

★★
make it available **as structured data** (e.g., Excel instead of image scan of a table)

★★★
make it available **in a non-proprietary open format** (e.g., CSV instead of Excel)

★★★★
use URIs to denote things, so that people can point at your stuff

★★★★★
link your data to other data to provide context



Source: https://5stardata.info/en/

Consider: the **FAIR** guiding principles

Findable

- Describe data in a data repository
- Mint persistent identifiers (e.g. doi, handle, et al)

Accessible

- Consider what can be published
- Ensure consent documentation
- Perform de-identification / anonymisation

Interoperable

- Use open formats
- Apply consistent vocabulary
- Use common/disciplinary metadata standards

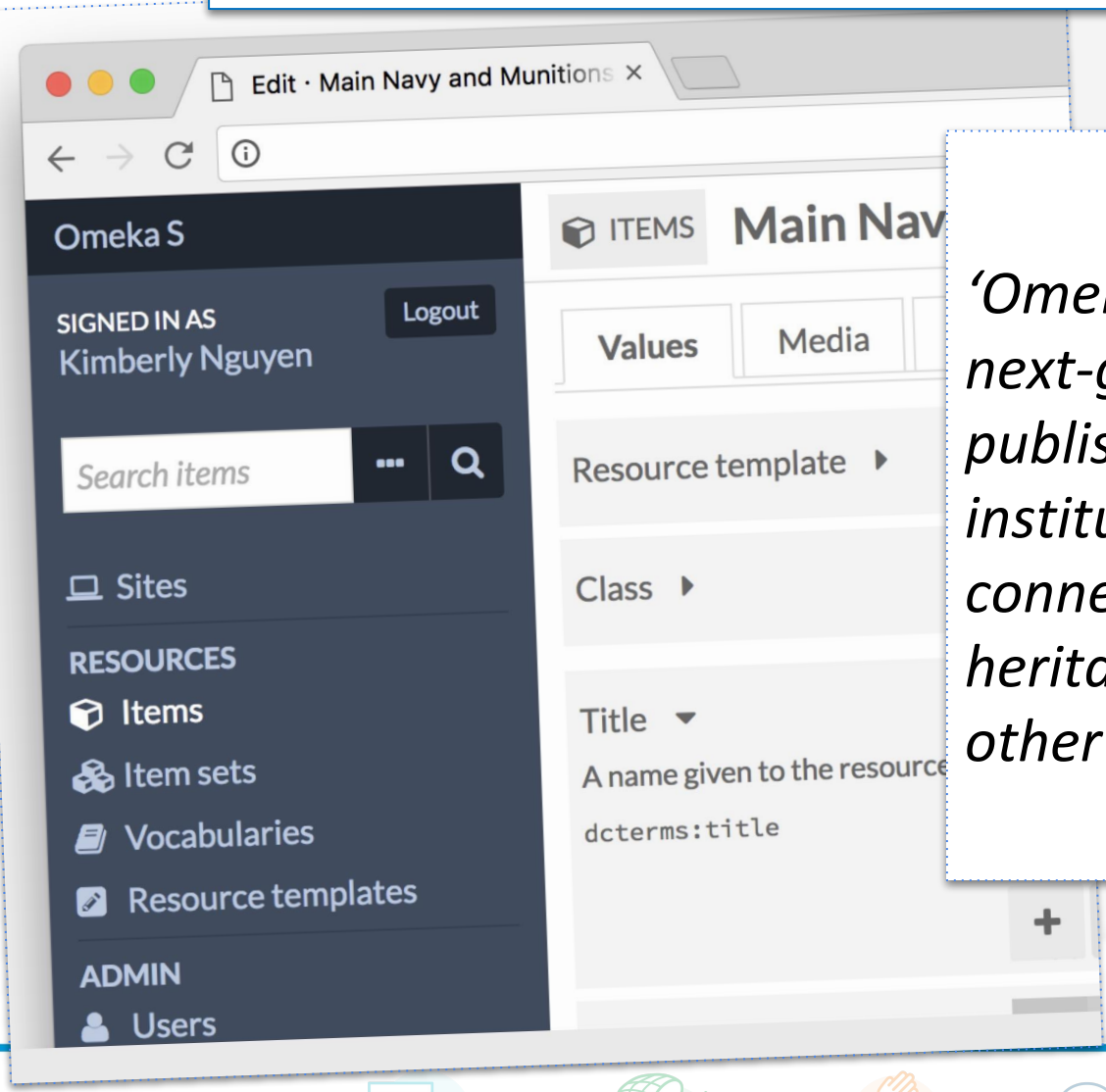
Reusable

- Consider permitted use
- Apply machine-readable open licenses (e.g. CC-BY etc.)

Adapted from: Zimmer, Niklas; King, Thomas (2018): **Data discovery and re-use**. figshare. Presentation. https://doi.org/10.25375/uct.7358423.v1

Omeka S

https://omeka.org/s/

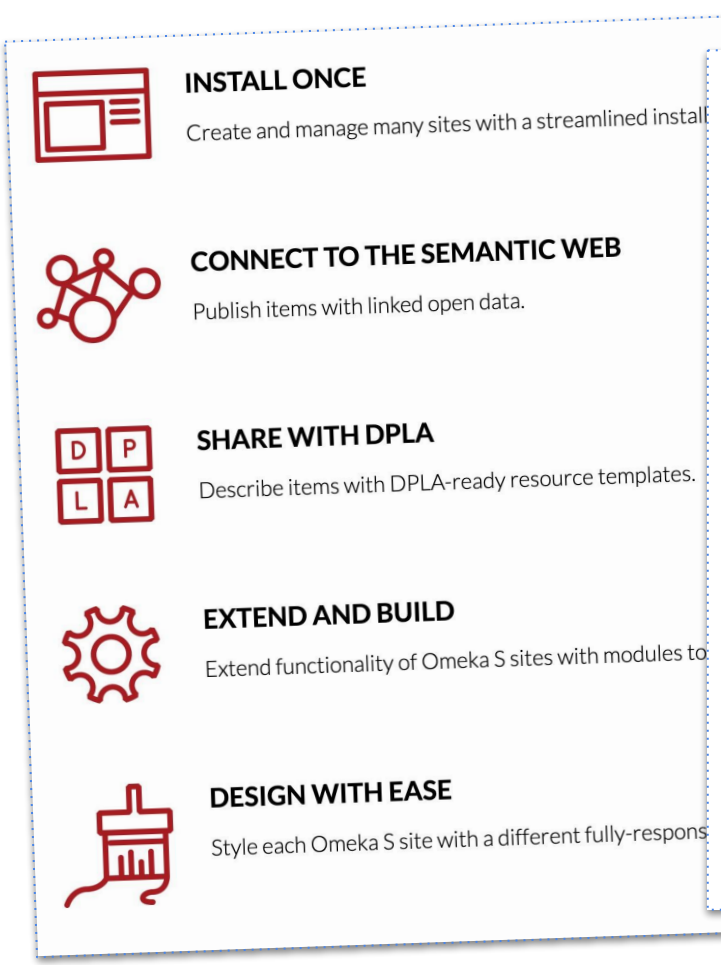


'Omeka S is a next-generation web publishing platform for institutions interested in connecting digital cultural heritage collections with other resources online.'

Source: https://omeka.org/s/

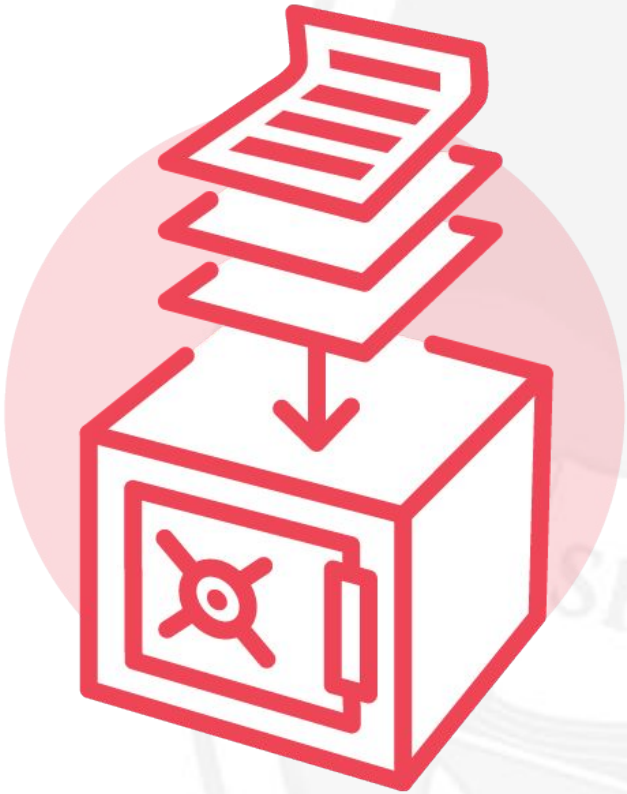
Omeka S & iiif server

https://omeka.org/s/



*'The **International Image Interoperability Framework** (IIIF) defines several application programming interfaces that provide a standardised method of describing and delivering images over the web, as well as "presentation based metadata"[1] (that is, structural metadata) about structured sequences of images. If institutions holding artworks, books, newspapers, manuscripts, maps, scrolls, single sheet collections, and archival materials provide IIIF endpoints for their content, any IIIF-compliant viewer or application can consume and display both the images and their structural and presentation metadata.'*

Source: https://en.wikipedia.org/wiki/International_Image_Interoperability_Framework



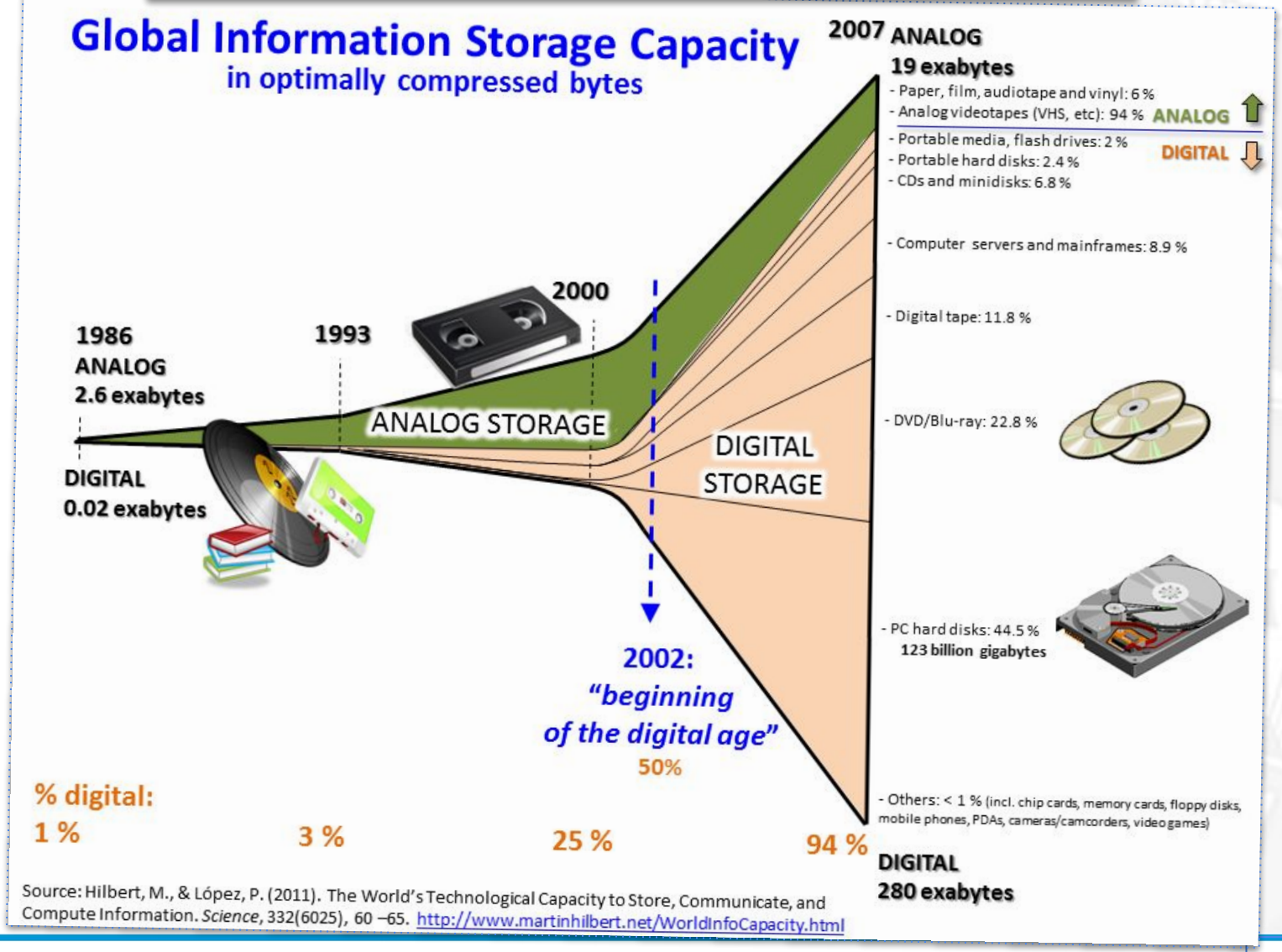
MANAGE, STORE, PRESERVE

Manage, Store, Preserve

Data needs to be findable, accessible, interoperable and reusable (FAIR), which is largely a question of machine-readable metadata. Only well-curated data can become a meaningful future resource.

- ★ create metadata and documentation
- ★ migrate data to best format
- ★ migrate data to suitable medium
- ★ back-up and store data
- ★ archive data

Keeping in mind the 'data deluge'



Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60 –65. http://www.martinhilbert.net/WorldInfoCapacity.html

Storage + Backup ≠ Preservation

Maintaining backups of stored data does not mean that they are digitally preserved. Digital preservation is an institutional endeavour to ensure that data remain accessible and usable in the long term, in view of:

- **technological change** (e.g. legacy media & formats)
- **bit-rot** (decay of digital files over time, e.g. on flash drives)
- **link-rot** (decay of identifiers over time, e.g. on websites)
- **media failure** (e.g. 'head crash' on hard drives, CD-Rs oxidising)

Digital preservation is handled by specialist staff (archivists, librarians), using dedicated hard- and software solutions. NB: Some data legally require digital preservation, something which is ideally planned for from the outset (see **DMP**).

Example: Digitisation for digital preservation

'legacy' media →	hardware →	software →	digital files			→ !
			master (preservation)	service (working)	access (access)	
Documents: manuscripts; theses; ...	flatbed scanner; feeder scanner; …	Acrobat Pro; …	.tif	.jpg	.jpg .txt .pdf	
Images: photographic prints; positives (slides); negatives; maps; ...	virtual drum scanner; digital camera & lighting equipment; map scanner, ...	Silverfast Studio; Nextimage; Photoshop; Lightroom; ...	.tif .fff .dng	.jpg	.jpg .png	
Audio: ¼-inch reel-to-reel; cassette; DAT; MD; ...	reel-to-reel, cassette, DAT and MD recorders; DAC; mixer; ...	Logic Pro; Waves Restoration Suite plugins; MediaHuman Audio converter; ...	.aif	.wav	.mp3	
Video: Umatic; Betacam; VHS; MiniDV; ...	Umatic, Betacam and VHS cleaners and recorders; MiniDV, DVCam descks; …	MediaExpress; FinalCut Studio; Premiere Pro; …	.mov .mpg2	.avi	.mp4	

Digitisation for Digital Preservation & Access

e.g.: http://www.digitalservices.lib.uct.ac.za/dls/what-we-digitise

A future problem: where are our data?

We know *where* they are but…

- It's in an unsupported file format
- It's in a legacy system
- It's not well described so it's irretrievable
- It's corrupted

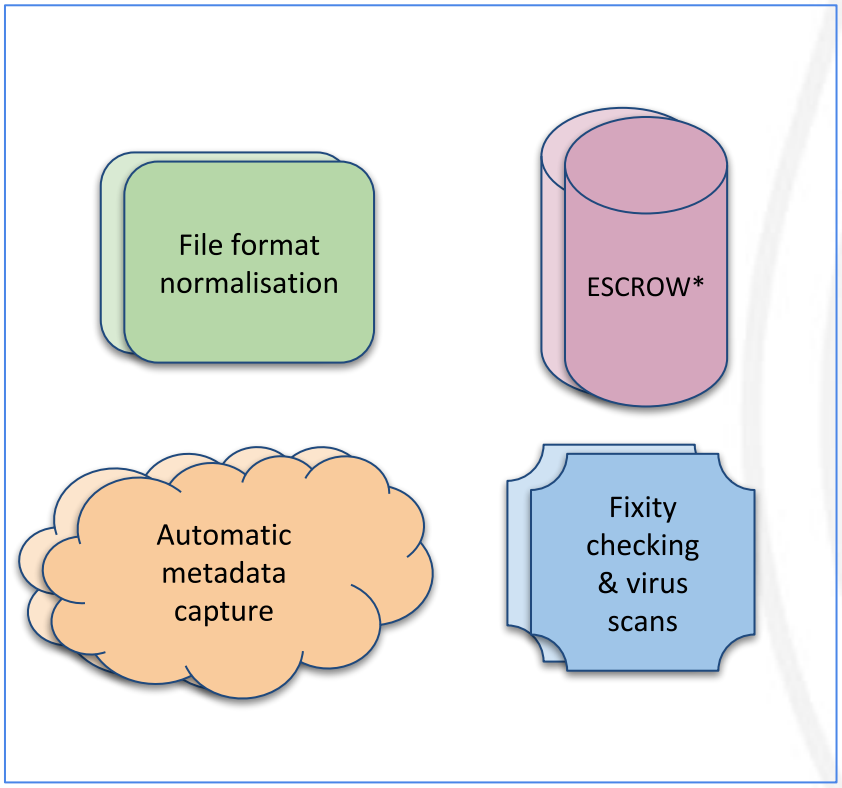
We *don't even know* where they are …

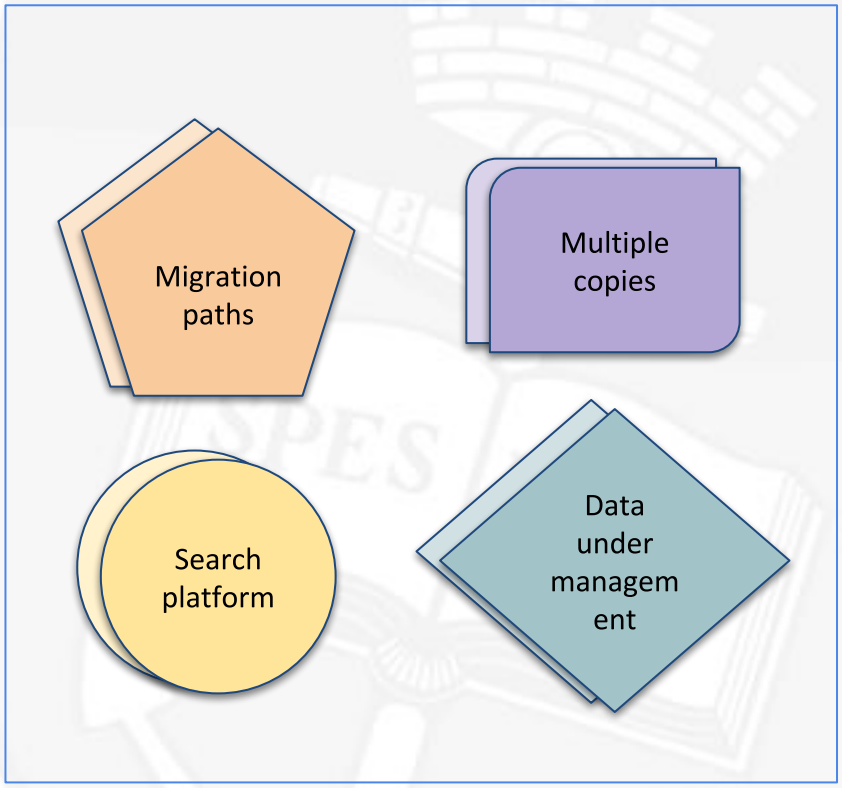
- It was on destroyed hardware
- A third party has it
- It's on a hard drive in a vault
- I expected it to be just where I left it

!

Adapted from: Arkivum: **Webinar Recording - Making the case for digital preservation.** Available: http://sites.arkivum.com/webinar-recording-making-the-case-for-digital-preservation-how-to-engage-your-internal-stakeholders-20-sept?hsCtaTracking=afd562aa-7fef-4f16-a1de-0958a8d68dce%7C277de3d6-6467-4c10-a387-8931548403fe

Digital preservation technologies & processes





** **'Source code escrow** is the deposit of the source code of software with a third-party escrow agent. Escrow is typically requested by a party licensing software (the licensee), to ensure maintenance of the software instead of abandonment or orphaning.'* Online. Available: https://en.wikipedia.org/wiki/Source_code_escrow

Adapted from: Arkivum: **Webinar Recording - Making the case for digital preservation.** Available: http://sites.arkivum.com/webinar-recording-making-the-case-for-digital-preservation-how-to-engage-your-internal-stakeholders-20-sept?hsCtaTracking=afd562aa-7fef-4f16-a1de-0958a8d68dce%7C277de3d6-6467-4c10-a387-8931548403fe

World Digital Preservation Day
7 November 2019

DIGITAL LIBRARY SERVICES

Digital Preservation: an all-inclusive conversation

- *DPC | World Digital Preservation Day (07.11.2019)*
- *Australasia Preserves | A digital preservation CoP*
- *NeDICC | Network of Data and Information Curation Communities*
- *RDM at UCT Slack Workspace & DLS Training*

Digital Preservation Coalition

Strategic Plan 2018 - 2022

- ***Community Engagement:*** *enabling a growing number of agencies and individuals in all sectors and in all countries to participate in a dynamic and mutually supportive digital preservation community.*
- ***Advocacy:*** *campaigning for a political and institutional climate more responsive and better informed about the digital preservation challenge; raising awareness about the new opportunities that resilient digital assets create.*
- ***Workforce Development:*** *providing opportunities for our members to acquire, develop and retain competent and responsive workforces that are ready to address the challenges of digital preservation.*
- ***Capacity Building:*** *supporting and assuring our members in the delivery and maintenance of high quality and sustainable digital preservation services through knowledge exchange, technology watch, research and development.*
- ***Good Practice and Standards:*** *identifying and developing good practice and standards that make digital preservation achievable, supporting efforts to ensure services are tightly matched to shifting requirements.*
- ***Management and Governance:*** *ensuring the DPC is a sustainable, competent organization focussed on member needs, providing a robust and trusted platform for collaboration within and beyond the Coalition.*

World Digital Preservation Day

Home > Events > World Digital Preservation Day

World Digital Preservation Day

7 November 2019 | 00:00 - 24:00 Worldwide

Advocacy

Dünya Sayısal Koruma Günü

At-Risk Digital Materials

About World Digital Preservation Day

World Digital Preservation Day is held on the first Thursday of every November. This year we will celebrate all things digital preservation on 7th November 2019!

#WDPD2019 on Twitter

Tim Gollins @timgollins

RT @Sarah_DPC: With the INTERNATIONAL Digital Preservation Conference #iPRES2019 in Amsterdam this week ... what better time to invite #dig...

VIEW TWEET

Sep 20

Tim Gollins @timgollins

Australasia Preserves

A digital preservation community of practice

DIGITAL LIBRARY SERVICES

AUSTRALASIA PRESERVES

A Digital Preservation Community Of Practice

HOME ABOUT CONTACT CODE OF CONDUCT JOIN BLOG POSTS

Australasia Preserves is a digital preservation community of practice (CoP) for Australia, New Zealand, and the broader Australasian region.

A **community of practice** is a group of people who share a craft or a profession

The **Australasia Preserves community of practice** is a group of people who share an interest in the field of digital preservation

Being part of a community of practice provides opportunities to develop personally and professionally

Australasia Preserves helps build local and global relationships, enabling sharing of digital preservation knowledge and experience

'South Africa Preserves' [?]

See: ***Australasia Preserves - a digital preservation community of practice***

DIGITAL LIBRARY SERVICES

Australasia Preserves *is a digital preservation community of practice for the Australasian region, with members from a wide range of domains across Australia and New Zealand. Community activities cover varied interests and areas of working with digitised and born-digital materials, encouraging participation and input from a wide variety of people and organisations, including libraries, archives, records, IT, government, and universities.*

[...] This community of practitioners, managers, educators, students, and enthusiasts enables sharing of digital preservation knowledge and expertise, and fosters opportunities to develop personally and professionally through an active online forum, monthly virtual meet-ups, locally organised events and networking opportunities, and working groups.

The community welcomes anyone interested in learning and sharing good practice for ongoing care of and access for digital materials to get involved, to connect, and to collaborate, in order to enhance digital preservation capability for a wide range of people and organisations in the Australasian region.

- *2018 highlights for the community: https://bit.ly/2FJrSUb*
- *The Australasia Preserves 2019 Briefing Pack: https://bit.ly/2GM82dD*
- *Join the online forum: https://groups.google.com/forum/#!forum/australasia-preserves*

Source: '***Australasia preserves***' (2019). (Online) Accessible: https://www.australasiapreserves.org/p/about.html

NeDICC

Network of **Data** and **Information Curation Communities**

NeDICC aims to promote the development and use of research data and information curation standards and practices to ensure the long term preservation and accessibility of digital research outputs in support of e-Research.

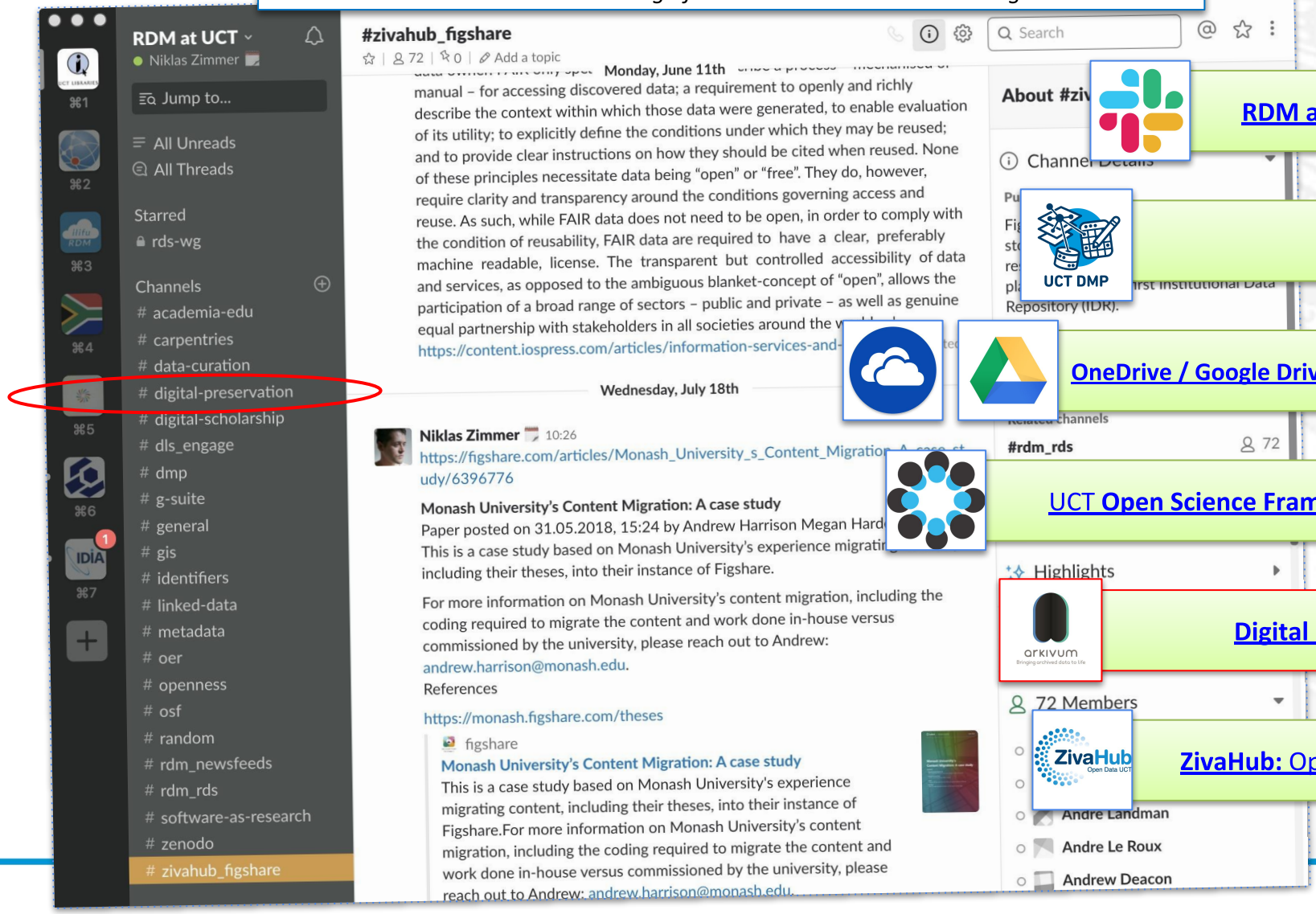
Specific aims and functions of the Network include:

- The provision of a forum, for practitioners and managers involved in digital object management practices, to exchange experience and express alternative views.
- Activities aimed at promoting communication and co-operation between members of the Network include meetings, seminars, workshops and conferences to:
  - address issues of interest/concern.
  - expose the community to new developments and trends, provide opportunities to engage with a wider audience, as well as showcase work and initiatives.
  - develop the knowledge and skills of members.
  - promote awareness and best practices relating to digital preservation, dissemination and use of research outputs.
  - collaborate on projects in support of shared objectives.

Source: https://nedicc.com/about-nedicc/

‘RDM at UCT’ Slack workspace

SLACK: ‘*Searchable Log of All Conversation and Knowledge*’



[RDM at UCT (Slack)]

[UCT DMP]

[OneDrive / Google Drive / MS Teams]

[UCT Open Science Framework (OSF)]

[Digital preservation]

[ZivaHub: Open Data UCT]

World Digital Preservation Day
7 November 2019

DIGITAL LIBRARY SERVICES

dls@uct.ac.za

@DigitalUct

rdm-at-uct.slack.com

DIGITAL LIBRARY SERVICES

http://www.digitalservices.lib.uct.ac.za/

UCT DMP

omeka s

arkivum
Bringing archived data to life

ZivaHub
Open Data UCT

FREY

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Thursday, 27th September 2019

World Digital Preservation Day
7 November 2019

DIGITAL LIBRARY SERVICES

THANK YOU

Some resources drawn on (i)

- DCC Curation Lifecycle Model by the Digital Curation Centre (DCC).
- Digital Curation Manual by the Digital Curation Centre (DCC).
- Digital Preservation Education for NC State Government Employees: Resources for general users.
- Digital Preservation for Beginners: Personal blog created by Megan Armaral, who received her MLis from San Jose State University, and became intrigued by the topic of digital preservation.
- Digital Preservation Handbook: The DPC's guide to the subject of managing digital resources over time and the issues in sustaining access to them.
- Digital Preservation Tutorial: Cornell/ICPSR Tutorial.
- Digital Tape Preservation Strategy: Preserving Data or Video? An essay about the philosophies and strategies applied to large scale video collections that are both born-digital and tape-based.
- JISC Beginner's Guide to Digital Preservation Blog: Guide to digital preservation (UK JISC projects).
- Library and Archives Canada: Strategy for a digital preservation program.
- Levels of Digital Preservation: from the National Digital Stewardship Alliance (NDSA).
- National Digital Information Infrastructure and Preservation Program: (NDIIPP) (Library of Congress).
- NEDCC Preservation Toolkit: Website documenting Surveying Digital Preservation Readiness: Toolkit for Cultural Organizations.
- OAIS Reference Model: introduction to the Open Archival Information System Reference Model.
- Parsimonious Preservation: Paper by Tim Gollins on the small, simple steps you can take to move forward in digital preservation.
- Personal Digital Memories Poster: Printable version of a poster presented at iPRES2010.

Some resources drawn on (ii)

- "Preserving Digital Culture": Slideshow about digital heritage and the importance of digital preservation.
- PRODUCING A BUSINESS CASE FOR DIGITAL PRESERVATION AND LONG-TERM ACCESS. Arkivum (2018).
- Take the quiz: Did you know that digital materials can be more difficult to preserve than physical materials?
- Team Digital Preservation: Videos created by Digital Preservation Europe about digital preservation challenges.
- The Digital Beyond: A blog about your digital existence and what happens to it after your death. An important source for archival, cultural, legal and technical insights to help you predict and plan for the future of your data.
- The Fragility of Digital Materials: Excerpt from Digital History: A Guide to Gathering, Preserving, and Presenting The Past on The Web that discusses in detail why digital formats are fragile and at-risk.
- The Library of Congress: Information on promoting personal archiving to individuals.
- The Signal: (Blog) From NDIIPP at the Library of Congress, this blog features frequent posts on a range of digital preservation and library technology topics.
- Webinar on Preserving Your Personal Digital Memories: (Windows Movie File) Archived ALCTS webinar from April 2011 that provides and overview to methods and tools for preserving personal digital information.
- "Why Digital Preservation is Important for Everyone": Short video on how digital content – unlike content on traditional media – depends on technology to make it available.
- Web archiving guidance.
- Why Digital Preservation is Important for Everyone: Library of Congress video, part of the NDIIPP Digital Preservation Video Series, exploring the complex nature of the digital preservation problem, how digital content depends on technology to make it available and requires active management to ensure its ongoing accessibility.
- You've Got to Walk Before You Can Run: first steps for managing born-digital content received on physical media, intended for anyone who does not know where to begin managing born-digital materials.