



Doing Digital Scholarship with Digital Library Services (DLS)

The Motivations

Wednesday, 15th May 2019, 10:00 - 11:00
Ulwazi training room

DLS Team

Digital Library Services

- [Niklas Zimmer](#)
- [Saniin Muftić](#)
- [Patricia Chikuni](#)
- [Ya'qub Ebrahim](#)
- [Thomas Slingsby](#)

ORCID

Connecting Research
and Researchers

Let's change
what we value
in research.



Sign
DORA





Digital Scholarship and DLS | **mission & vision**

We provide **open**, online access to primary resources for teaching, learning and research at the University of Cape Town (UCT) through digitisation, **digital scholarship**, data curation and preservation services.

We subscribe to and support the practice of **Open Science**.

Source: DLS website: <http://www.digitalservices.lib.uct.ac.za/>

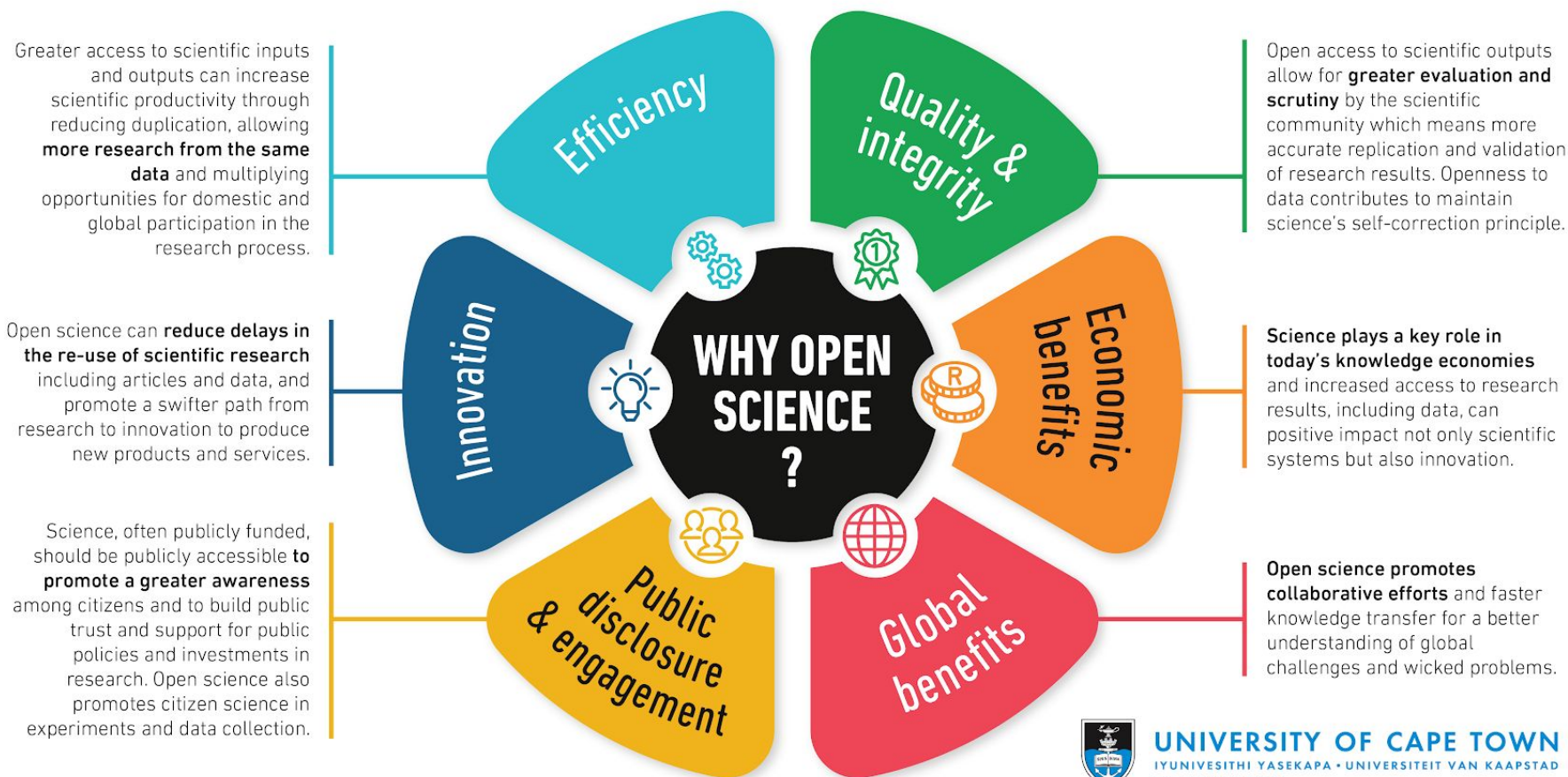
Open Science

Open Science is the movement to make scientific *research* (including publications, data, physical samples, and software) and its *dissemination* **accessible to all levels** of an inquiring society, amateur or professional.

Open Science is arguably simply proper science. Others are enabled to **collaborate and contribute**, since research data [...] and other research processes are **freely available**, under terms that enable **reuse, redistribution and reproduction** of the research and its underlying data and methods and subscribe to grounded ethical practices.

Source: Foster Open Science: (<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>)
Adapted from: Woelfle, M.; Olliaro, P.; Todd, M. H. (2011). "Open science is a research accelerator". Nature Chemistry. 3 (10): 745–748. <https://doi.org/10.1038%2Fncchem.1149>

Open Science at UCT



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD
GRAPHICS BY GAELIN PINNOCK

Source: UCT RDM Why Open Science: https://commons.wikimedia.org/wiki/File:UCT_RDM_Why-Open-Science.png

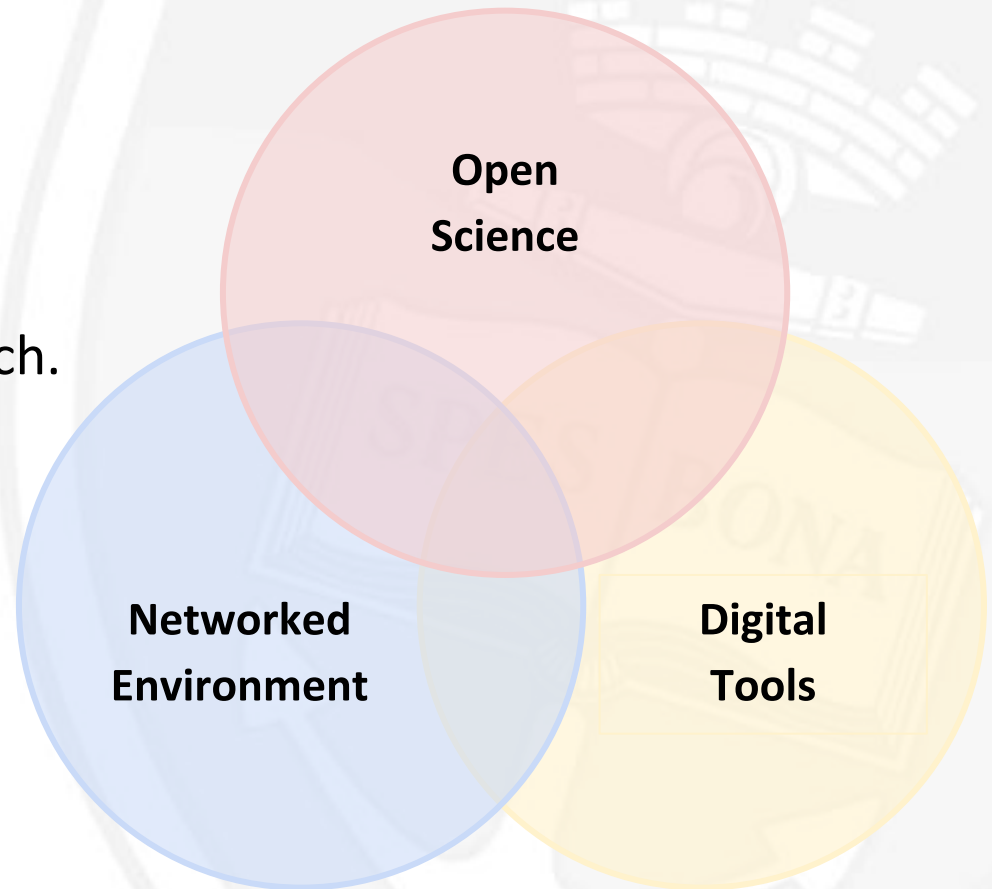


Digital Scholarship and Research Data Management *a brief overview*



What is **Digital Scholarship**?

Digital Scholarship is the application and integration of digital tools and methods to discover, *research* and teach.



Source: Weller, M. 2011. *The Digital Scholar*; Adapted from: <https://www.open.edu/openlearn/ocw/mod/oucontent/view.php?id=48677§ion=2>

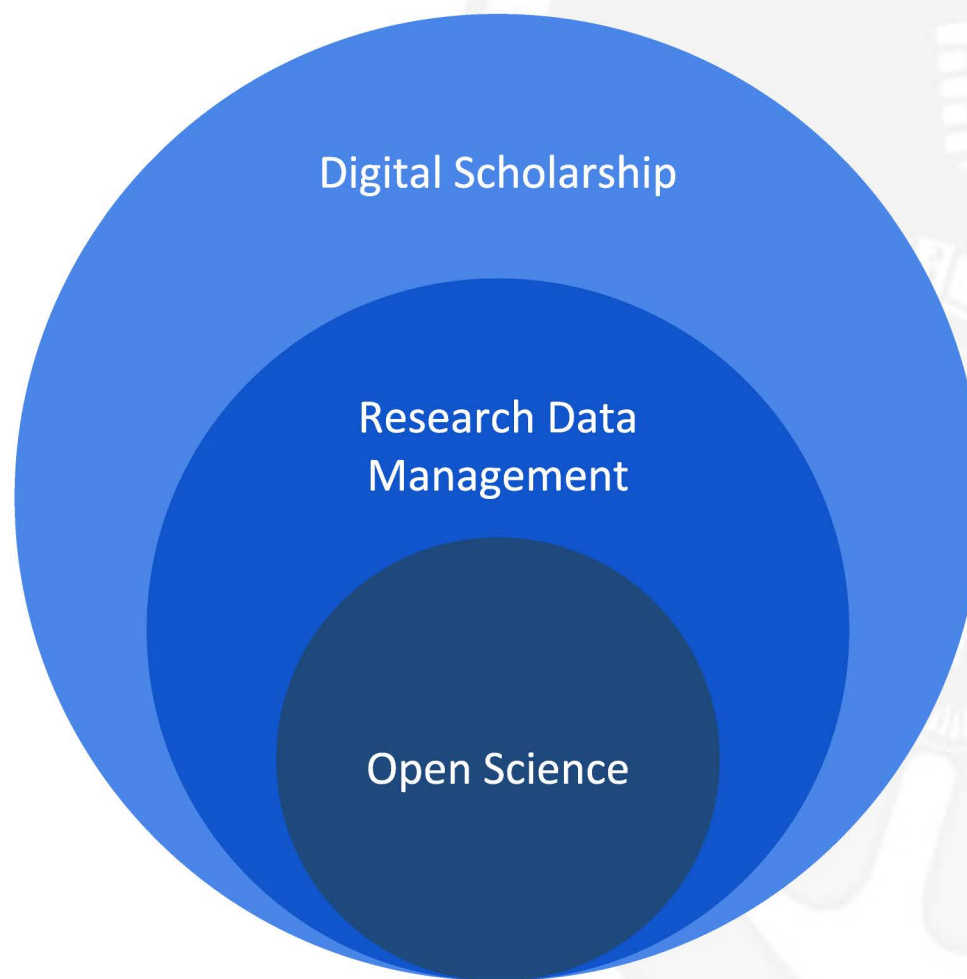
The research data management (RDM) lifecycle



What is Research data management (**RDM**)?

- The **organisation and documentation** of the data processes (collection, description, de-identification, curation, archiving and publication) within a research project.
- Already practised by researchers, but generally for internal use, and to varying degrees of professionalism.
- Part of an international drive towards **Open Science**, to professionalise data management practices, and make research more coherent and shareable.
- Good **Digital Scholarship** practices along every step of the research lifecycle.

DS - RDM - OS





What is (your) data
and why should you make it **reusable**?

“What (are my) research data?”

QUALITATIVE | QUANTITATIVE

RESEARCH DATA

Micro
Unit record
Raw
Field
Experimental
Spatial
Cleaned
Processed
Primary
Secondary
De-identified

Documents (text, spreadsheets)
Lab notebooks, field notebooks, diaries
Questionnaires, transcripts, surveys
Codebooks
Films, audio or video tapes/files
Photographs, image files
Sensor readings
Test responses
Artifacts, specimens, physical samples
Models, algorithms, scripts
Content analysis
Focus group recordings; interview notes

OBSERVATIONAL | EXPERIMENTAL | SIMULATION | DERIVED

Compiled from: LibGuides@ Macalester University. Available at: <https://libguides.macalester.edu/c.php?g=527786&p=3608583>



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE

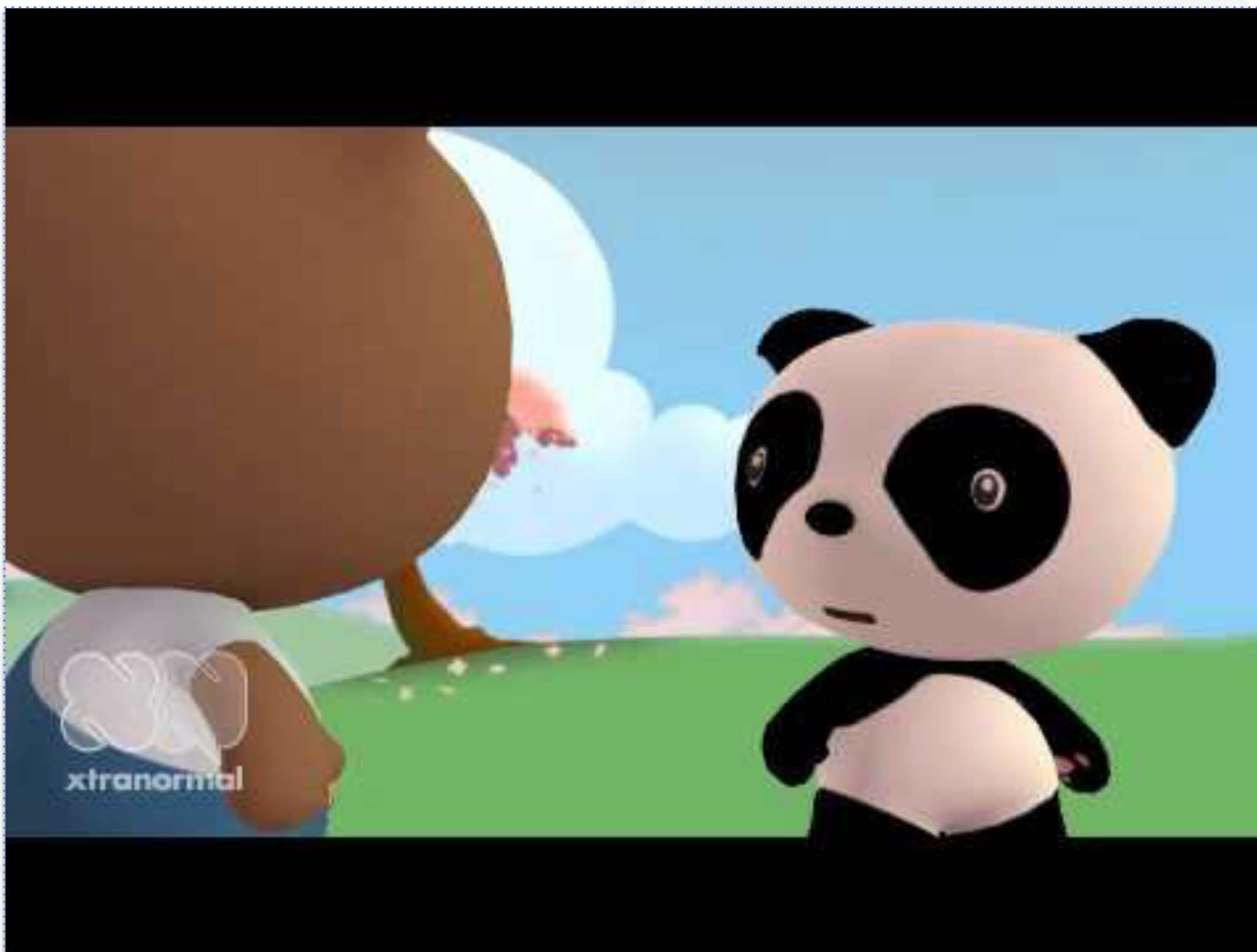


SHARE & PUBLISH



MANAGE, STORE, PRESERVE

What is **closed** science?



Source: NYU Health Sciences Libraries. <https://youtu.be/NzK3sAtr-4>

'Good RDM makes data reusable'

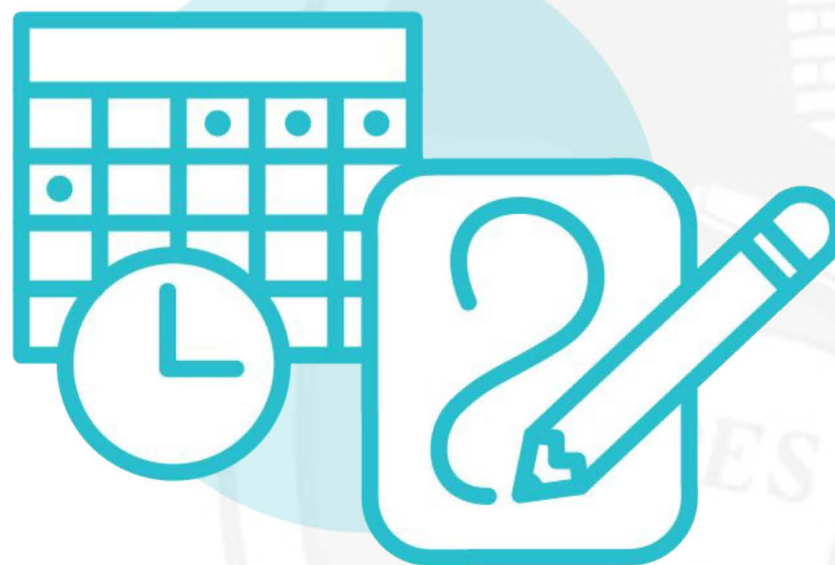


Source: [10 aspects of highly effective research data - Good research data management makes data reusable](#) By Anita de Waard, Helena Cousijn, PhD, and IJsbrand Jan Aalbersberg, PhD



The Research Data Lifecycle

*Digital Scholarship tools and methods to assist
with Research Data*



PLAN & DESIGN



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

Fail to plan ... plan to fail!



Source: Tomasz Sienicki, Cycling in Denmark (2009). Available: https://commons.wikimedia.org/wiki/File:Cyklisci_dk_ubt.JPG



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

What is a DMP & why create one?

A **data management plan** (DMP) is a living, written document explaining what you intend to do with your data during and following the conclusion of your research project.

A DMP is already a **requirement** by many **funders** (NIH, Wellcome Trust, NRF).

Even when it is not a requirement, having made such a plan can **save you time** and **effort** during your research, as it assists you with **organising your data**, preparing it for the next step in its lifecycle, and clarifying who will have access to it, how, and when.

A DMP provides **guidance for curation-specific activities**, such as file-naming, archiving, formats suitable for long-term preservation, etc.

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>



The new **student MoU** at UCT

Comply with **institutional requirements**: In 2019, a new **student MoU** (Memorandum of Understanding) was implemented for all postgraduate researchers, requiring them to create a DMP as part of the registration process:

★ **E.3 Research data management policy**

The requirement for storage of research data as specified by funders must be met - i.e. of both research and scholarship / bursaries. (See: <http://www.researchsupport.uct.ac.za/managing-research-data>)

The supervisor and candidate should confirm that they are aware of the requirement to complete and submit a Data Management Plan (DMP) (available on the Library website <http://www.digitalservices.lib.uct.ac.za/dls/rdm-planning>) prior to collecting, storing, describing or analysing data.

Confirm that this requirement has been complied with by indicating 'Yes' below.

Are you aware of the research data management policy?

Supervisor	<div style="border: 2px solid red; padding: 5px; display: inline-block;"> Yes <input type="checkbox"/> </div>
Student	<div style="border: 2px solid red; padding: 5px; display: inline-block;"> Yes <input type="checkbox"/> </div>

10 January 2019
Page 6
ACA47a

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>



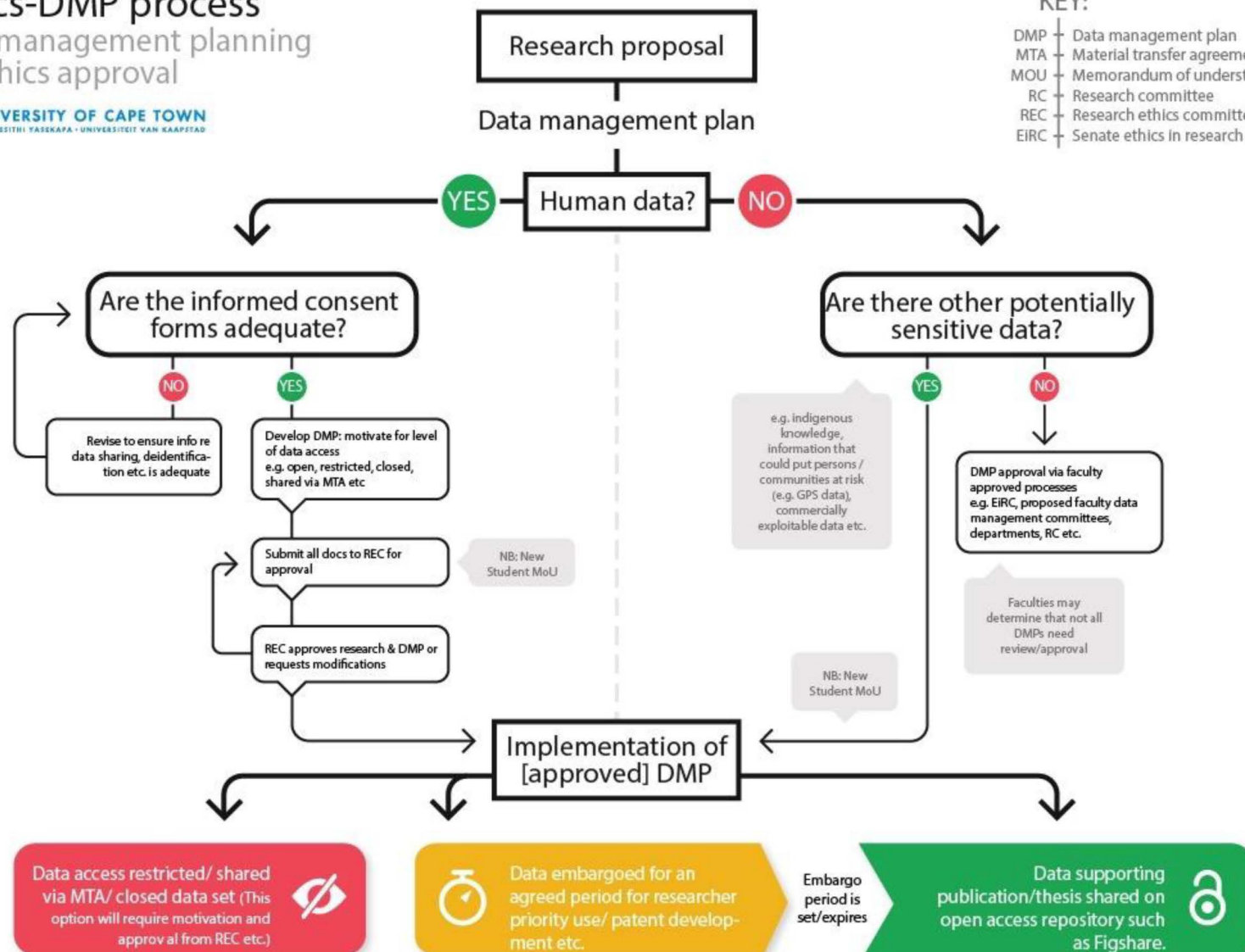
Ethics-DMP process

Data management planning for ethics approval



KEY:

- DMP + Data management plan
- MTA + Material transfer agreement
- MOU + Memorandum of understanding
- RC + Research committee
- REC + Research ethics committee
- EIRC + Senate ethics in research committee



accessible: https://commons.wikimedia.org/wiki/File:RDMGraphics_Ethics_final-Cropped.jpg



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

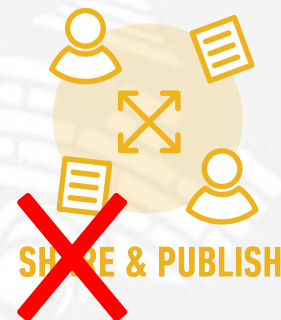
Simplified 'traditional' research process



PLAN & DESIGN



COLLECT & CAPTURE



SHARE & PUBLISH



COLLABORATE & ANALYSE



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

Simplified 'modern' research process

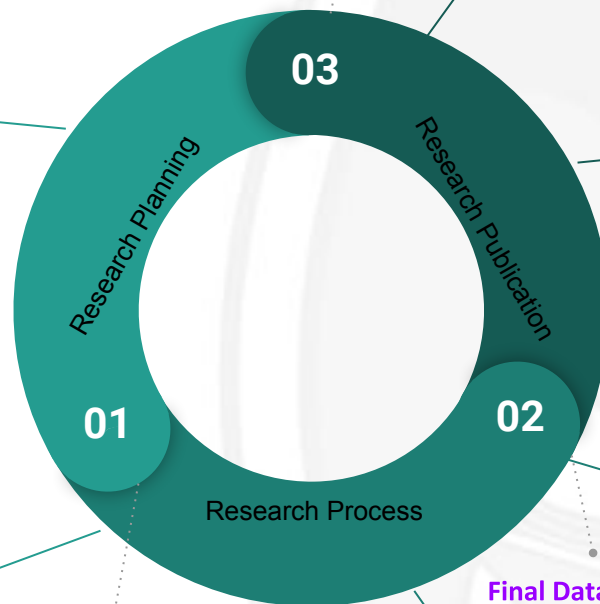


PLAN & DESIGN



COLLECT & CAPTURE

Initial Data Management Plan



Data Publication



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



COLLABORATE & ANALYSE

Final Data Management Plan



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH

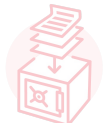


MANAGE, STORE, PRESERVE

Additional benefits of creating a DMP

- **Enhance your reputation:** Well documented, clean and organized data can lead to new collaboration and funding opportunities.
- If you plan on sharing your data, a DMP can help you work through the issues you need to address to make sharing possible.
- Finally, a DMP helps ensure that your data remains **useable** to yourself, your collaborators, and other researchers in future.

Adapted from: OSF Guides > Best Practices > Handling Data > Creating a data management plan (DMP). Available: <http://help.osf.io/m/bestpractices/1/618674-creating-a-data-management-plan-dmp>



My plan (Gender, Health and Justice Research Unit)

Plan details

GHJRU DMP

Share

Export

This page gives you an overview of your plan. It tells what your plan is based on and gives an overview of the questions that you will be asked.

Plan name	My plan (Gender, Health and Justice Research Unit)
ID	-
Grant number	-
Principal Investigator/Researcher	Ya'qub Ebrahim
Plan data contact	-
Description	-

This plan is based on:

Institution | University of Cape Town (UCT-Generi)

Pick from a variety of templates (funder-specific or generic, i.e. 'UCT') to assist you with planning how you will collect, store, manage and analyse your research data during your research project.

Sections	Questions
1. Project name	- Insert the name of your project proposal.
2 Introduction/type of study	- Provide a summary of the written description of the proposed study. Include the study's objectives, design, and methods.
3. Description of existing data	- Provide if possible a survey of previously existing data relevant to the project; the nature and scale of such data; and a brief discussion of whether and how these data will be integrated or the gaps in these datasets the new study will fill.
4. Data collection and generation	- TYPES OF DATA/DATA OUTPUTS - Describe what types of data will be collected. Indicate whether the data will be qualitative or quantitative and the likely file formats in which the data will be collected. Indicate if there is an intention to convert file formats for long-term accessibility and preservation. - METHODOLOGIES FOR DATA CREATION/GENERATION - Describe the how data will be collected for this study. - QUALITY MANAGEMENT - Describe the quality control (QC) measures and quality assurance (QA) measure you will implement.
5. Data management, documentation and curation	- MANAGING, STORING AND CURATING DATA - Indicate how you will be storing and curating your electronic and paper/hard copy data. Focus on principles and systems with brief examples, and avoid long lists. - DATA DOCUMENTATION - Indicate what additional documentation (aside from the DMP) if any will accompany the dataset to support future users. - FILE NAMING CONVENTIONS - Indicate the naming convention for your data files. - DATA ARCHIVING - Outline your plans for storage/archiving of the final datasets. - ETHICS AND PRIVACY - Indicate how informed consent will be handled in your project.

DMPonline

<https://dmp.lib.uct.ac.za/>

My plan (Gender, Health and Justice Research Unit)

0/18 questions answered

approx. 12% of available space used

Plan details **GHJRU DMP** Share Export

1. Project name (1 question, 0 answered)

2. Introduction/type of study (1 question, 0 answered)

3. Description of existing data (1 question, 0 answered)

4. Data collection and generation (3 questions, 0 answered)

TYPES OF DATA/DATA OUTPUTS - Describe what types of data will be collected. Indicate whether the data will be qualitative or quantitative and the likely file formats in which the data will be collected. Indicate if there is an intention to convert file formats for long-term accessibility and preservation.

B I [List Icon] [Table Icon] [Link Icon] [Grid Icon]

Useful information is provided at every step.

Save

Not answered yet

METHODOLOGIES FOR DATA CREATION/GENERATION - Describe the how data will be collected for this study.

B I [List Icon] [Table Icon] [Link Icon] [Grid Icon]

Guidance Add comment

UCT Guidance

Data collected and stored by the GHJRU typically includes the following:

- In-depth interview audio files (mp3) and transcripts (MS word documents)
- Focus group discussion audio files (mp3) and transcripts (MS word documents, Nvivo files)
- Notes from in-depth interviews and focus group discussions, and other fieldnotes (MS word documents, Nvivo files)
- Quantitative survey data: both electronic (CSV, STATA, SPSS) and paper
- Minutes of research meetings—to be considered "data" only if collected as the result of a research process (Microsoft word documents)

Accessibility and preservation

Open and machine-readable formats help preserve data in the long term. Consider converting text files into RTF, PDF or XML format, quantitative data into CSV, and audio files into WAV to ensure they are accessible for future users and software systems.

Guidance Add comment

UCT Guidance



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

Typical DMP questions

- **What type of data** will be generated in your research?
- How will your data be **named and referenced**?
- What **file formats** are involved?
- What data and **metadata standards** will you follow?
- Who will **have access** to your data?
- How and when will you **share** your data, if applicable?
- Will you be **digitally preserving** your data? If yes, how so?
- How will you **license** your datasets?
- How will you ensure **privacy** or **confidentiality**, if applicable?

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>



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



COLLECT & CAPTURE



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

Advice for the Collect & Capture Process

While collecting and capturing your data, make sure that you document it with correct, meaningful **metadata**:

- Describe the type of data generated:
 - The **form** (*What kind of data does it hold?*)
 - The **stability** of each dataset (*How does it change over time?*)
 - Create **unique names** for each of your datasets
- Document the data you are capturing, and how you are identifying it within each data set by building a **data dictionary**.
- Practice good **file naming conventions**.
- **Document your process** and store it together with your data (e.g. readme.txt).

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>



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH

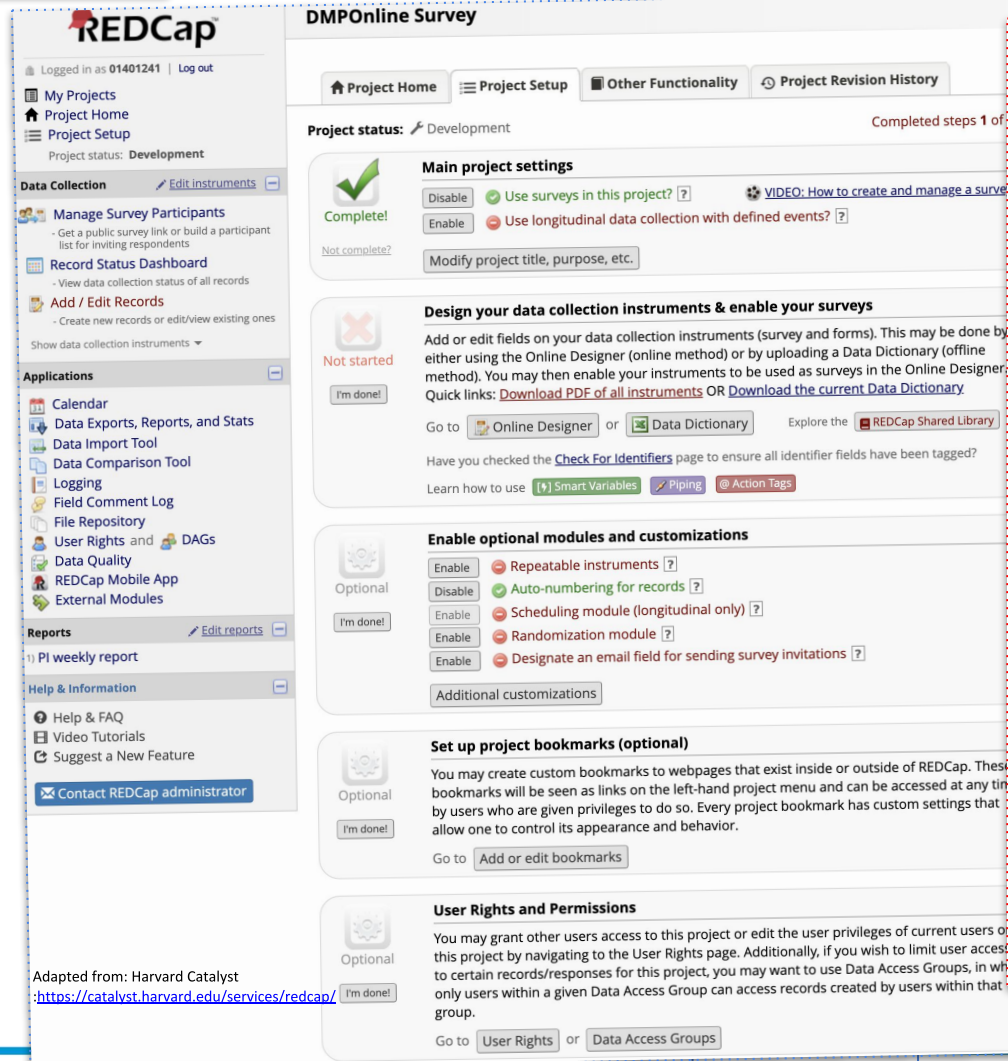


MANAGE, STORE, PRESERVE



RedCap

<https://trn-redcap.uct.ac.za/>



REDCap
Logged in as 01401241 | Log out

My Projects
Project Home
Project Setup
Project status: Development

Data Collection [Edit instruments](#)

Manage Survey Participants
Get a public survey link or build a participant list for inviting respondents

Record Status Dashboard
View data collection status of all records

Add / Edit Records
Create new records or edit/view existing ones

Show data collection instruments

Applications

- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging
- Field Comment Log
- File Repository
- User Rights and DAGs
- Data Quality
- REDCap Mobile App
- External Modules

Reports [Edit reports](#)

1) PI weekly report

Help & Information

- Help & FAQ
- Video Tutorials
- Suggest a New Feature

[Contact REDCap administrator](#)

DMPOnline Survey

[Project Home](#) [Project Setup](#) [Other Functionality](#) [Project Revision History](#)

Project status: Development Completed steps 1 of 2

Main project settings

Complete!

Disable ☒ Use surveys in this project? [VIDEO: How to create and manage a survey](#)

Enable ☐ Use longitudinal data collection with defined events? [?](#)

Not complete? [Modify project title, purpose, etc.](#)

Design your data collection instruments & enable your surveys

Not started

[I'm done!](#)

Add or edit fields on your data collection instruments (survey and forms). This may be done by either using the Online Designer (online method) or by uploading a Data Dictionary (offline method). You may then enable your instruments to be used as surveys in the Online Designer. Quick links: [Download PDF of all instruments](#) OR [Download the current Data Dictionary](#)

Go to [Online Designer](#) or [Data Dictionary](#) Explore the [REDCap Shared Library](#)

Have you checked the [Check For Identifiers](#) page to ensure all identifier fields have been tagged?

Learn how to use [Smart Variables](#) [Piping](#) [Action Tags](#)

Enable optional modules and customizations

Optional

[I'm done!](#)

Enable ☐ Repeatable instruments [?](#)

Disable ☒ Auto-numbering for records [?](#)

Enable ☐ Scheduling module (longitudinal only) [?](#)

Enable ☐ Randomization module [?](#)

Enable ☐ Designate an email field for sending survey invitations [?](#)

[Additional customizations](#)

Set up project bookmarks (optional)

Optional

[I'm done!](#)

You may create custom bookmarks to webpages that exist inside or outside of REDCap. These bookmarks will be seen as links on the left-hand project menu and can be accessed at any time by users who are given privileges to do so. Every project bookmark has custom settings that allow one to control its appearance and behavior.

Go to [Add or edit bookmarks](#)

User Rights and Permissions

Optional

[I'm done!](#)

You may grant other users access to this project or edit the user privileges of current users of this project by navigating to the User Rights page. Additionally, if you wish to limit user access to certain records/responses for this project, you may want to use Data Access Groups, in which only users within a given Data Access Group can access records created by users within that group.

Go to [User Rights](#) or [Data Access Groups](#)

A secure web application for building and managing online surveys and databases, useful for collecting and tracking information and data from research studies, scheduling study events and conducting surveys.

Features:

- input data from anywhere in the world
- projects can be used by researchers from multiple sites and institutions
- total control of shaping your database or survey
- data may be imported from external data sources to begin a study or to provide mid-study data uploads
- export survey results to common data analysis packages
- generate a PDF version for printing in order to collect responses offline

Adapted from: Harvard Catalyst
<https://catalyst.harvard.edu/services/redcap/>





COLLABORATE & ANALYSE



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



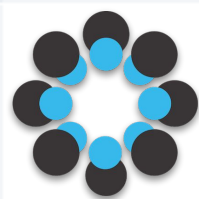
DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



Open Science Framework (OSF)

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

Research Methods PRACTICAL in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring

Contributors: Tamas Nagy, Zoltan Kekes

Date created: 2019-02-11 01:24 AM | Last Updated: 2019-04-30 02:46 PM

Category: Project

Wiki

Practical slides can be found here:

<https://drive.google.com/drive/folders/1brpFv87IO1f1Uye6zyad9jYSajocFca77usp=sharing>

Files

Name	Modified
Research Methods PRACTICAL in Clinical and Health Psychology - PS...	
- Dropbox: Readings and lecture slides to OSF	
+ Lecture slides to OSF	
+ Mini-exam questions and results	
+ readings	
- Google Drive: slides	
Practical 1 - Managing research projects, introducing OSF.gslides	2019-03-18 12:41 PM
Practical 12 - Writing an abstract.gslides	2019-04-29 09:38 AM
Practical 2 - Creating online questionnaires.gslides	2019-02-18 04:06 AM
Practical 3 - Reading, writing, and citing research papers.gslides	2019-02-25 02:14 PM
Practical 4 - Ethical issues in conducting and publishing resear...	2019-03-18 03:48 AM
Practical 5 - Intervention studies and group design.gslides	2019-03-18 12:42 PM
Project evaluation rubric.gsheet	2019-04-28 09:18 PM
- OSF Storage (Germany - Frankfurt)	

Citation

Recent Activity

- Tamas Nagy updated file Research Methods in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring
- Tamas Nagy updated file Research Methods in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring
- Tamas Nagy updated wiki page Home to version 2 of Research Methods in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring
- Tamas Nagy updated wiki page Home to version 1 of Research Methods in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring
- Zoltan Kekes linked Dropbox folder to Research Methods in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring
- Zoltan Kekes authorized the Dropbox addon for Research Methods in Clinical and Health Psychology - PSYM17-CH-107 - 2019 Spring

- free, online platform that allows you to register your project, manage stakeholders, and centralise data that might be stored at different locations with different collaborators
- 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

Show rows with cells including:

Variable	Variable name	Mesasurement unit	Allowed values	Description
Participant ID number	ID	Numeric	001-999	ID number assigned to participant in sequential order
Group number	GROUP	Numeric	1-30	Group assigned to participant based on ID number
Age in years	AGE	Numeric	18.0-65.0	Age of participant in years
Date of birth	DOB	mm/dd/yyyy	1-12/1-31/1951-1998	Participant's date of birth
Gender	SEX	Numeric	1 = male 2 = female	Participant's gender
Date of survey	SURVEY	mm/dd/yyyy	01/01/2015 - 01/01/2016	When the participant completed the survey
Self-reported consumer spending	SPEND	Numeric	0-100,000,000	Self-reported average yearly expenditure
Market sentiment	SENTIMENT	Numeric	1 = negative 2 = neutral 3 = positive	Sentiment towards US domestic economy
Actual GDP growth	GDP	Numeric	-5.0-5.0	Average US yearly GDP growth



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE

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.

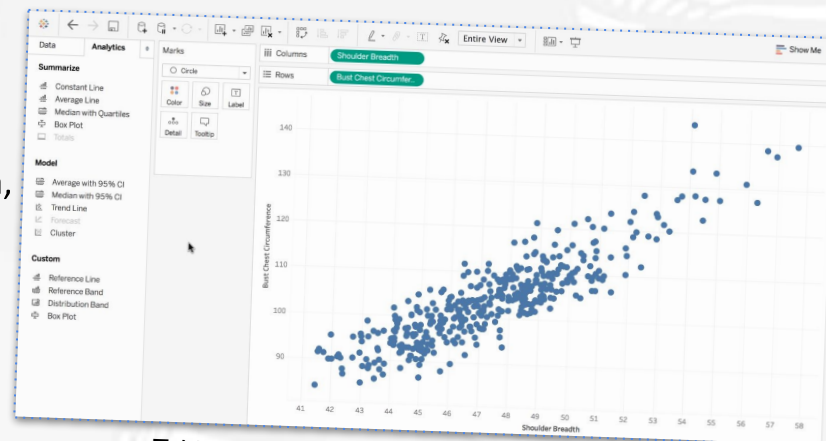


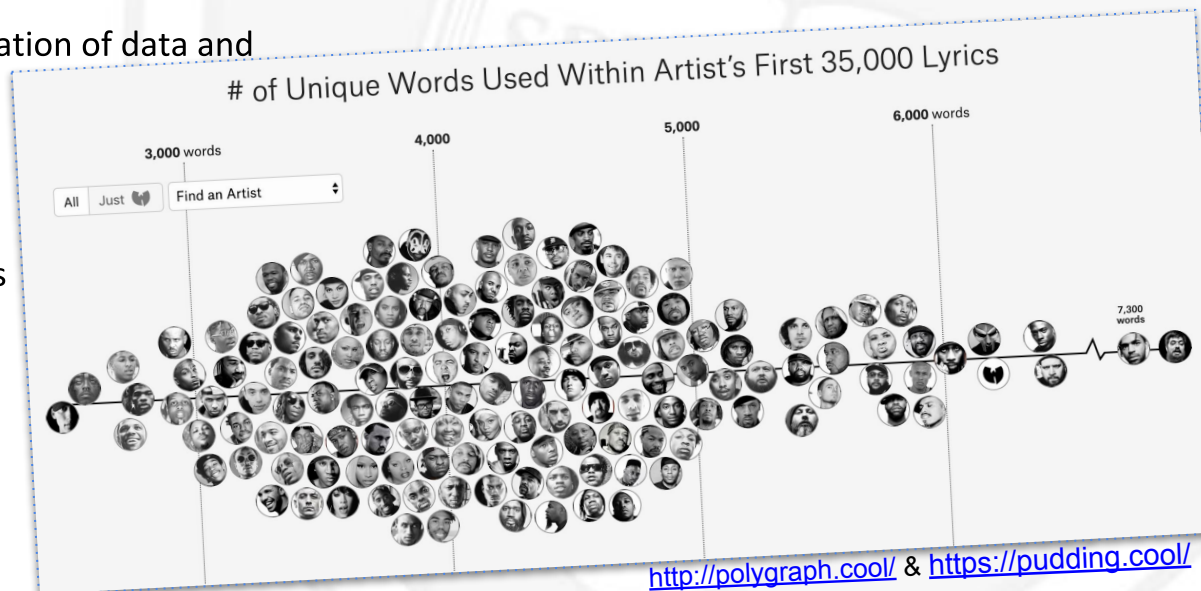
Tableau - <https://www.tableau.com/>

Data Visualization:

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

Digital Humanities:

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



<http://polygraph.cool/> & <https://pudding.cool/>

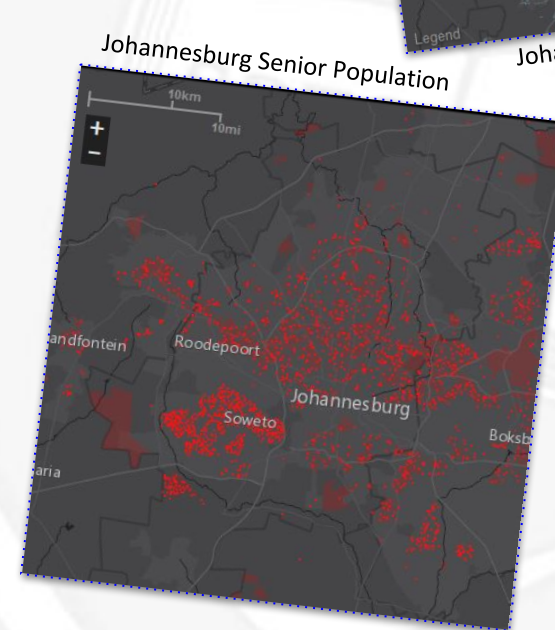
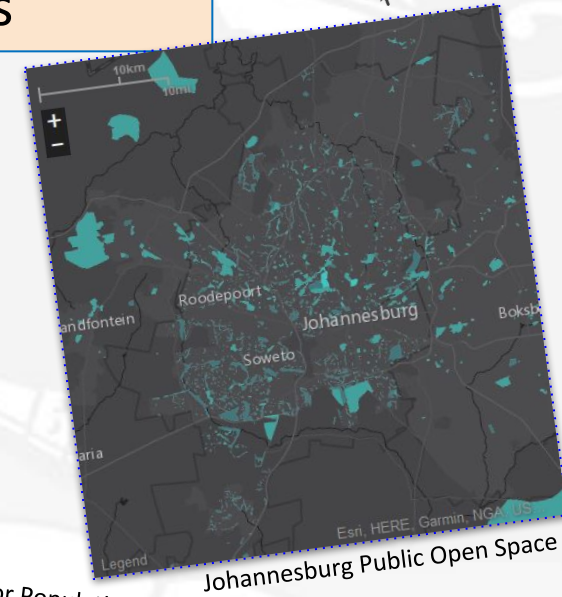
Geographic Information Systems

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.

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.



Images sourced from [UrbanObservatory.org's App](https://urbanobservatory.org/)

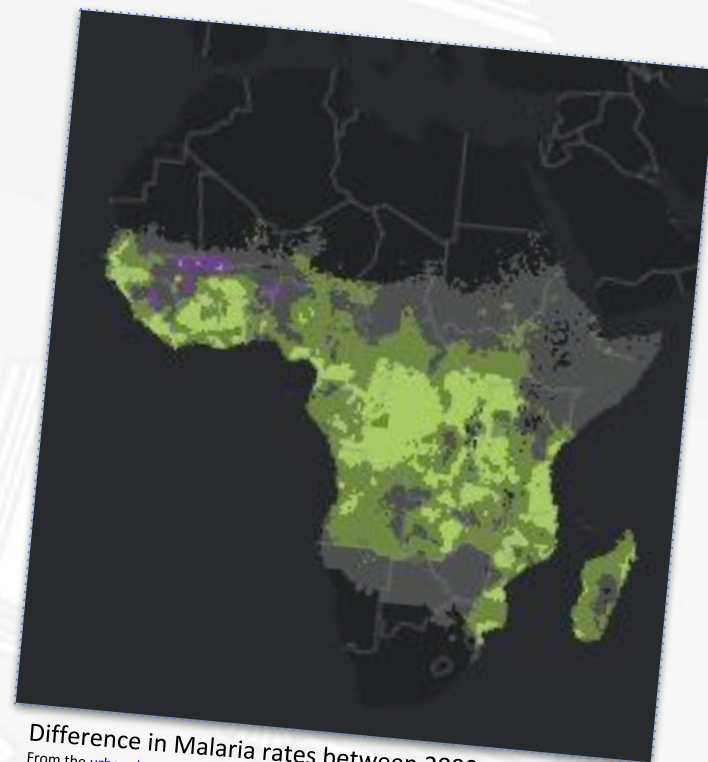
Data Analysis (Making Information):

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.

DLS’ GIS services assist with GIS software acquisition, project planning, troubleshooting, analysis and cartographic design.

Find us @ www.gis.uct.ac.za



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



DISCOVER, REUSE & CITE



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE

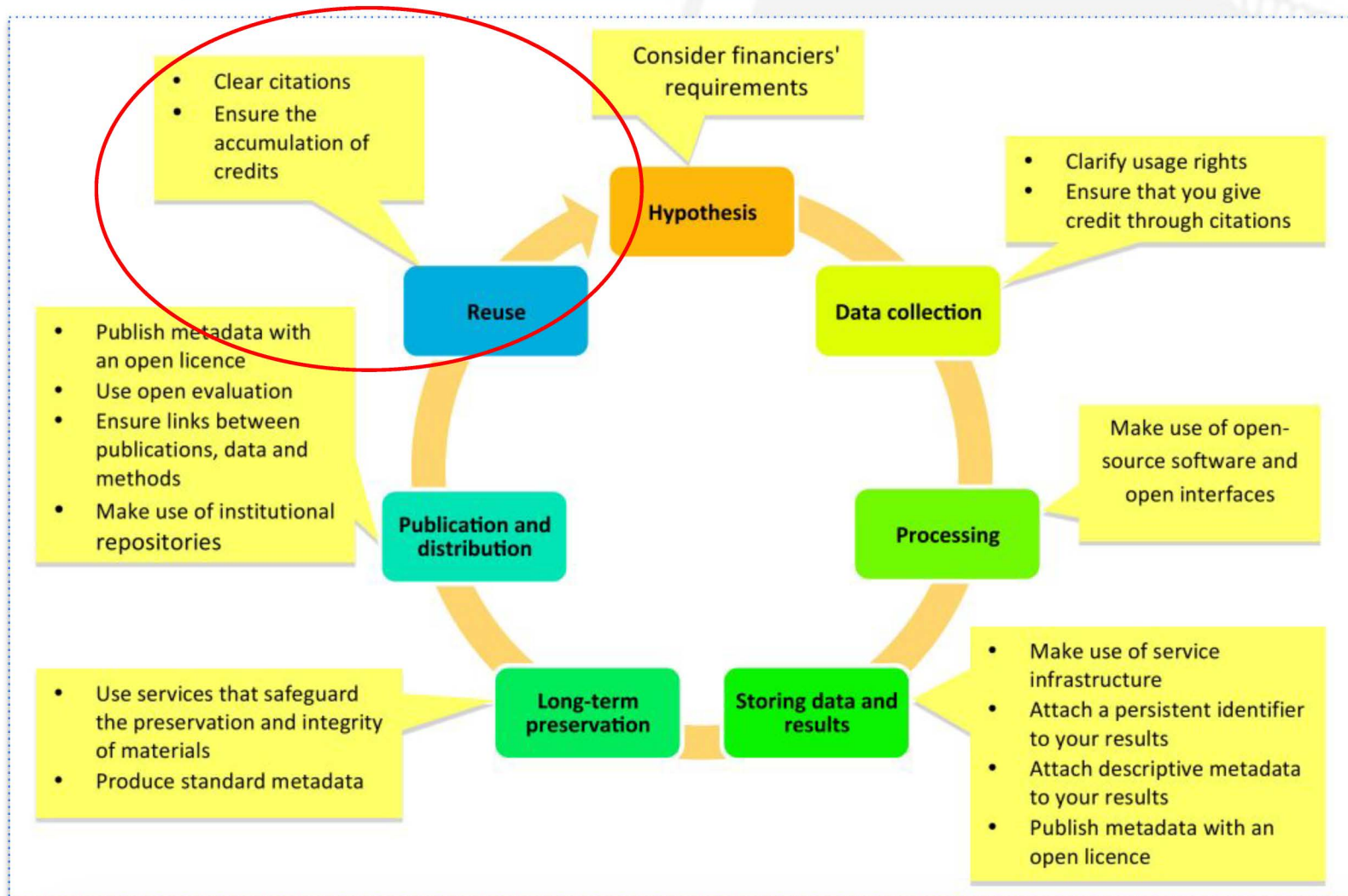


SHARE & PUBLISH



MANAGE, STORE, PRESERVE

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>

Two research data repositories at UCT

DataFirst and ZivaHub are registered, certified, and transparent, through independent review, standards, and policies.



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

A small overview of data catalogues, registries and repositories

directly UCT-relevant

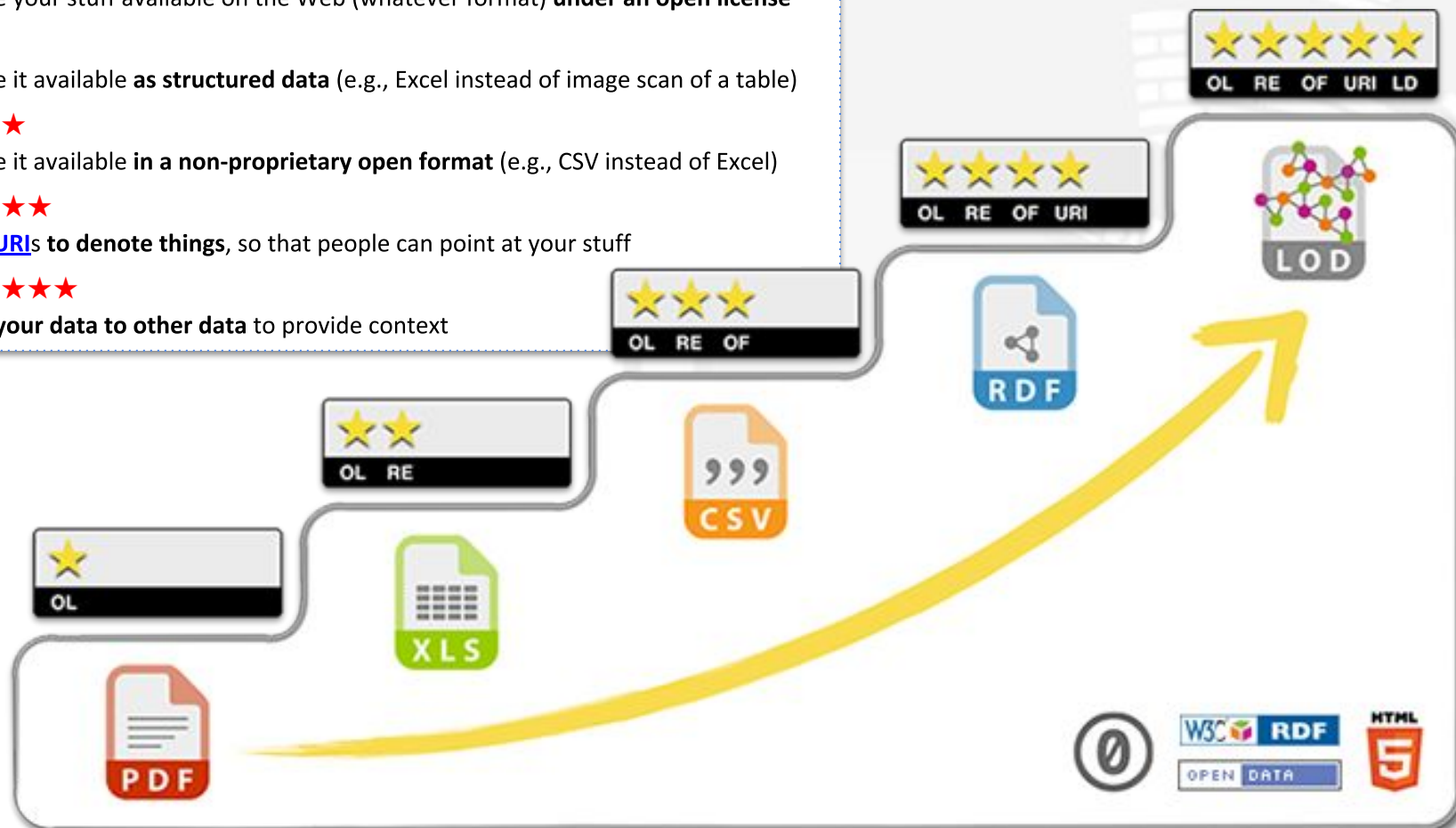
- [BioLINCC](#) – Clinical specimen database.
- [Dataverse](#) – Widely used open source repository system; Example: [HARVARD Dataverse](#)
- [dataMED](#) – prototype biomedical data search engine to discover data sets across data repositories or aggregators.
- [Code Ocean](#) – Cloud-based computational platform which provides a way to share, discover and run published code.
- [ContentMine](#) – Uses machines to liberate 100,000,000 facts from the scientific literature.
- [DataBank](#) – Analysis and visualisation tool that contains collections of time series data on a variety of topics.
- [DataCite](#) – Establish easier access to research data by providing persistent identifiers for data.
- [Datahub](#) – Publish or register datasets, create and manage groups and communities
- [Dataverse Network](#) – Harvard-based tool to share, cite, reuse and archive research data.
- [Deveo](#) – Free, private Git, Mercurial, and SVN repository management platform.
- [Dryad](#) – Data repository system for any files associated with any published article in the sciences or medicine.
- [Figshare\(.com\)](#) – Free cloud service for managing, sharing & publishing research data.
- [GenBank](#) – Gene sequence database provided by the National Center for Biotechnology Information.
- [GitHub](#) – Online software project hosting using the Git revision control system.
- [How Can I Share It](#) – Information and tools to ensure your articles can be shared with your colleagues easily.
- [Open Science Framework](#) – Open registration, version control & collaboration software system.
- [Quip](#) – Combines chat, documents, spreadsheets, checklist, and more to collaborate on any device.
- [re3data](#) – Global registry of research data repositories.
- [Research Compendia](#) – Tools for researchers to connect data, code & computational methods to published research.
- [SlideShare](#) – Community for sharing presentations and other professional content.
- [Zenodo](#) – A home for the long-tail of science, enabling researchers to share and preserve any research outputs.
- [ZivaHub | Open Data UCT](#) – UCT's digital repository.



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

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/>

Working with the **FAIR** guiding principles

- Describe your data in a data repository
- Receive a persistent identifiers (e.g. uct doi provided by ZivaHub)

Findable

- Consider what can be published
- Obtain participant consent
- Perform de-identification / anonymisation

Accessible

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>





SHARE & PUBLISH

What Stops you from sharing data?

1. **Misinterpretation** of the data.
2. **Misappropriation** of the data.
3. Damage to the researcher's **reputation** (CODATA-ICSTI, 2013).
4. **Myths** that scientific findings using shared data cannot be published in high impact journals (Milham, et.al 2018).

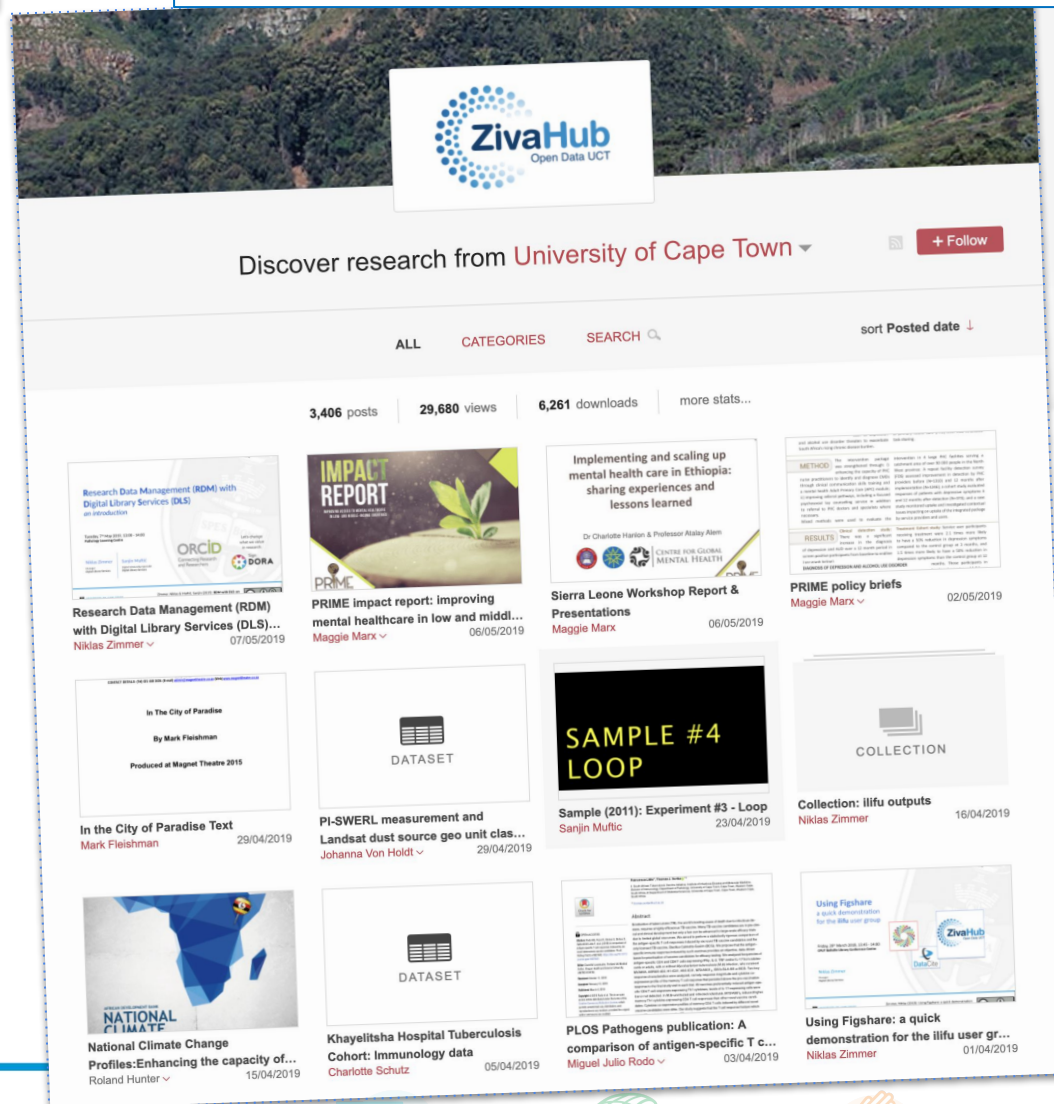
*However these fears immediately disappear the moment the data are properly managed and documented (CODATA-ICSTI, 2013).

Why share your research data?

1. **Funding agencies and institutions** now require the results of scientific studies to be shared with the public as a condition for providing grants or awards.
2. **Publishers** are now asking authors to deposit some datasets in public platforms.
3. To **confront some of the biases in data collection and analysis** (Atici et. al, 2013).
4. To **reproduce** or to **verify** research.
5. To enable others to **ask new questions** of extant data.
6. To **advance the state of research** and innovation (Borgman, 2012).
7. To **increase citation rate** (Piwowar, Day and Fridsma, 2007).
8. Sharing on platforms like Figshare **increases the visibility of individual researchers and their work online** (Peters et al., 2015).

ZivaHub | Open Data UCT

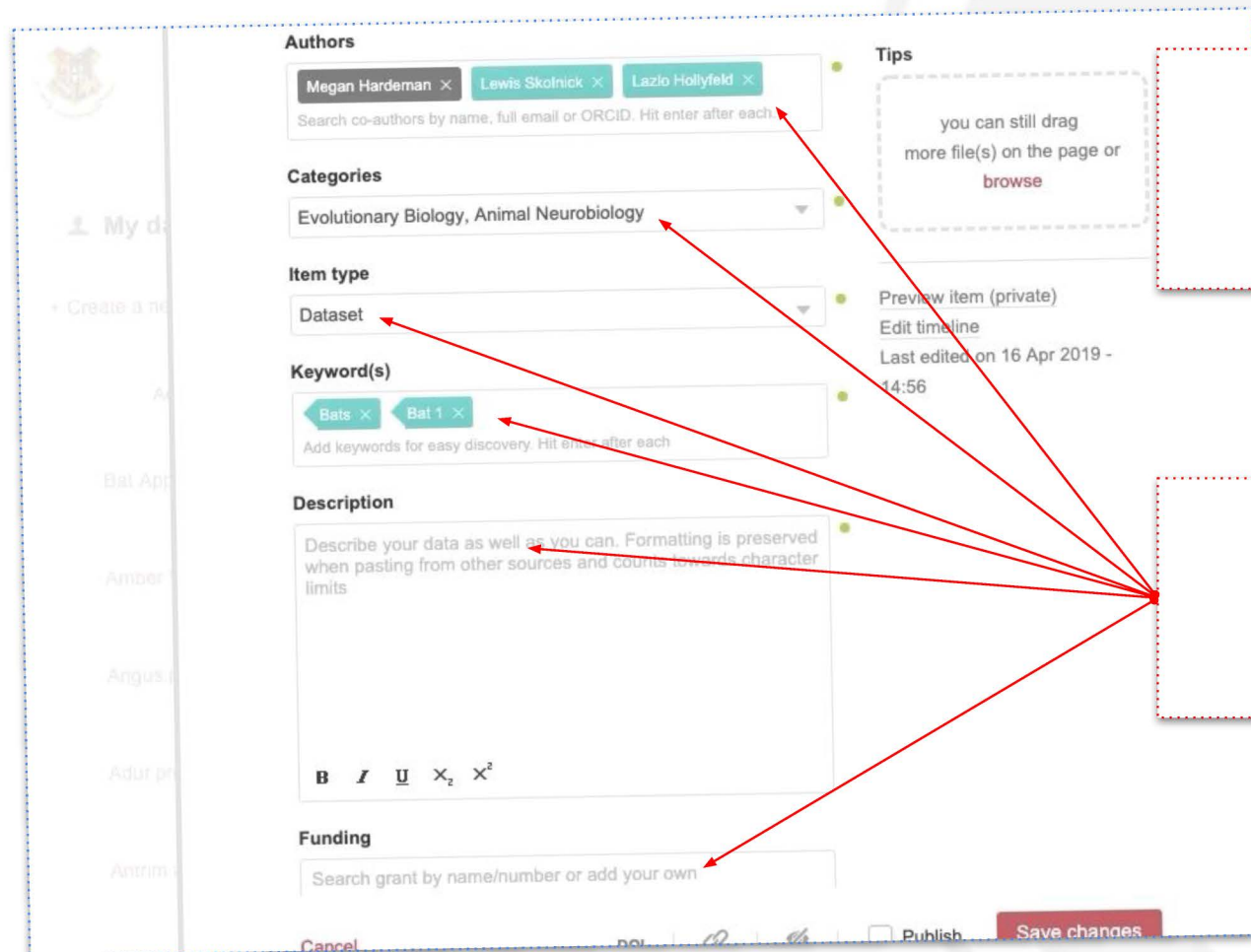
<https://zivahub.uct.ac.za/>



- a repository to store and openly disseminate data
- powered by Figshare for institutions
- keeps track of views, downloads and citations
- allows search across all Figshare platforms

ZivaHub | Open Data UCT

<https://zivahub.uct.ac.za/>



The screenshot shows the 'Create a new item' form on ZivaHub. Red arrows point from callout boxes to the following fields:

- Authors:** A list of authors (Megan Hardeman, Lewis Skolnick, Lazlo Hollyfeld) with a search prompt: 'Search co-authors by name, full email or ORCID. Hit enter after each.'
- Categories:** A dropdown menu showing 'Evolutionary Biology, Animal Neurobiology'.
- Item type:** A dropdown menu showing 'Dataset'.
- Keyword(s):** A list of keywords (Bats, Bat 1) with a search prompt: 'Add keywords for easy discovery. Hit enter after each.'
- Description:** A text area with a prompt: 'Describe your data as well as you can. Formatting is preserved when pasting from other sources and counts towards character limits.' Below the text area are formatting icons: Bold (B), Italic (I), Underline (U), Subscript (x₂), and Superscript (x²).
- Funding:** A text input field with a prompt: 'Search grant by name/number or add your own'.

On the right side of the form, there is a 'Tips' section with the text: 'you can still drag more file(s) on the page or browse'. Below this is a 'Preview item (private)' section with 'Edit timeline' and 'Last edited on 16 Apr 2019 - 14:56'.

At the bottom of the form are buttons for 'Cancel', 'Publish', and 'Save changes'.

supports the upload of **any file format**, and aims to visualise all of them

embeds relevant **metadata**, to make data **FAIR** compliant

Source: Figshare. End User Guide GIFs. Available: <https://figshare.com/s/b3600c85f576d88d067b>

ZivaHub | Open Data UCT

<https://zivahub.uct.ac.za/>

We track usage statistics, including views, downloads, citations, and Altmetrics. Citations are measured using ReadCube, a portfolio company of Digital Science.

B *I* U \times_2 \times^2

Funding

DCAT-AP for Wikibase and Wikidata

+ Add another grant

References

<https://phytopatholres.biomedcentral.com/articles/10.1186/s424>

Licence (what's this?)

- CC BY
- CC BY
- CC-0
- MIT
- GPL
- GPL-2.0
- GPL-3.0
- Apache-2.0

Cancel Publish Save changes

Tips

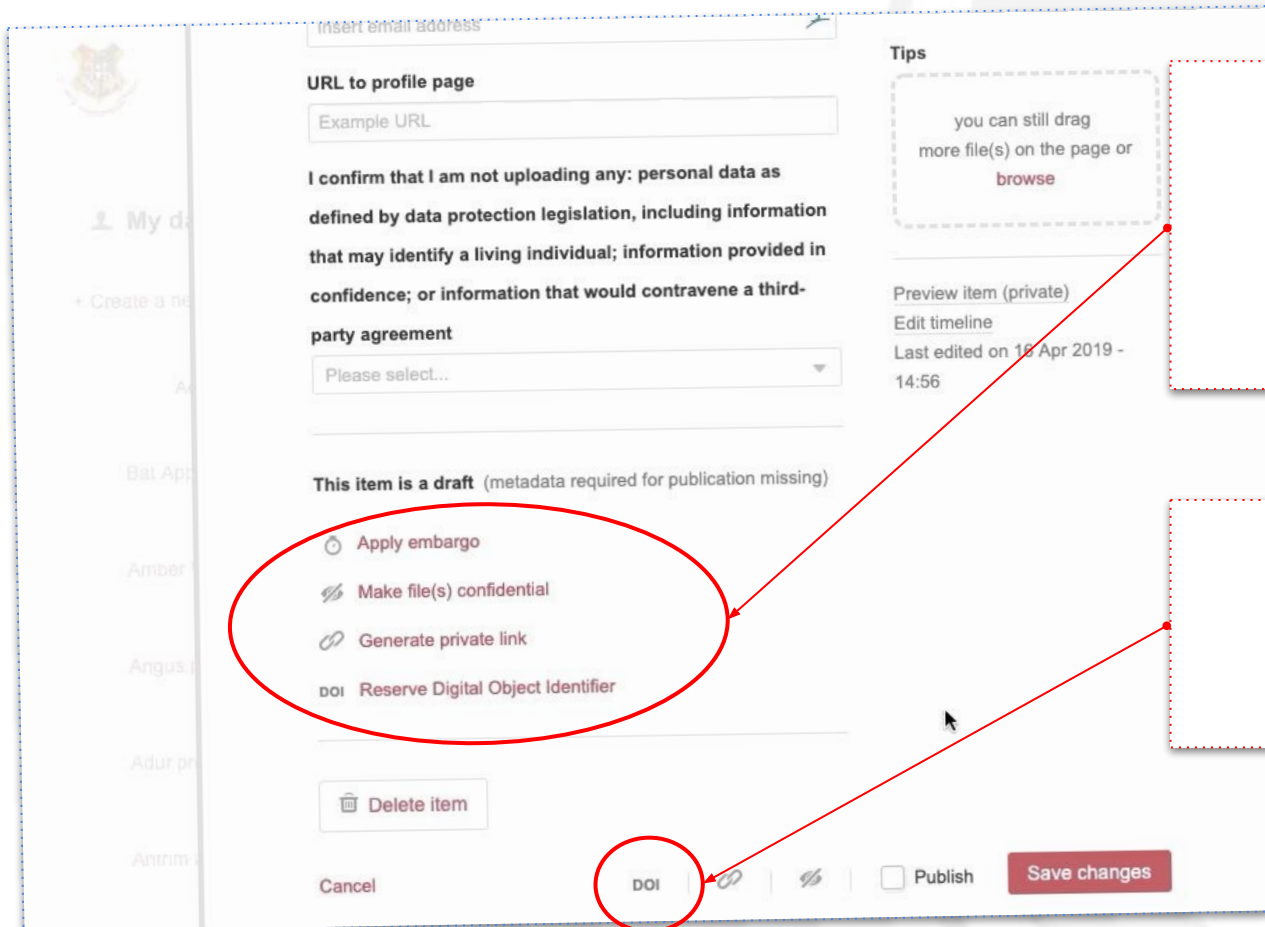
you can still drag
more file(s) on the page or
browse

choose from a range of
licensing options when
publishing your data
openly

Source: Figshare. End User Guide GIFs. Available: <https://figshare.com/s/b3600c85f576d88d067b>

ZivaHub | Open Data UCT

<https://zivahub.uct.ac.za/>



insert email address

URL to profile page
Example URL

I confirm that I am not uploading any: personal data as defined by data protection legislation, including information that may identify a living individual; information provided in confidence; or information that would contravene a third-party agreement

Please select...

This item is a draft (metadata required for publication missing)

- ☐ Apply embargo
- ☐ Make file(s) confidential
- ☐ Generate private link
- ☐ DOI Reserve Digital Object Identifier

Preview item (private)
Edit timeline
Last edited on 16 Apr 2019 - 14:56

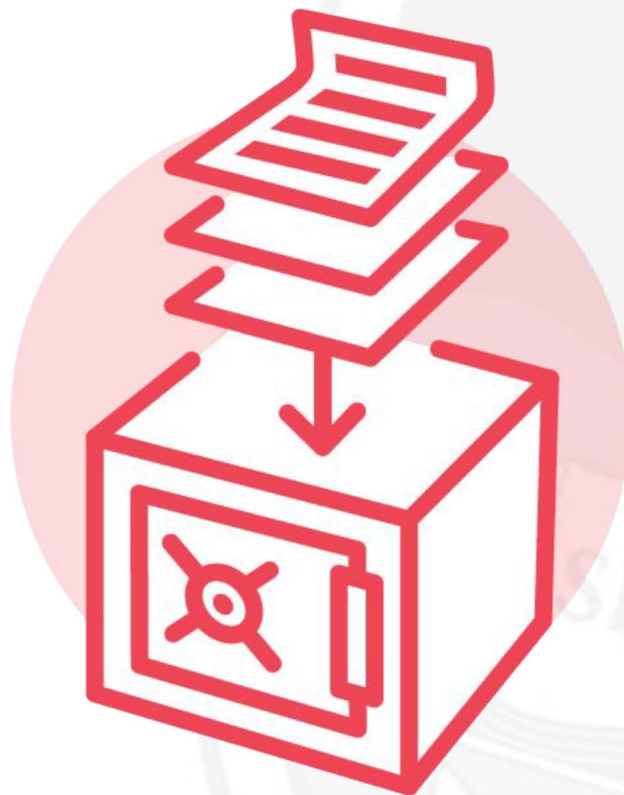
Delete item

Cancel DOI ☐ Publish Save changes

share data privately with your funders, reviewers, or supervisors before publishing it later

all item records get assigned a **Digital Object Identifier (DOI)**

Source: Figshare. End User Guide GIFs. Available: <https://figshare.com/s/b3600c85f576d88d067b>



MANAGE, STORE, PRESERVE



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE

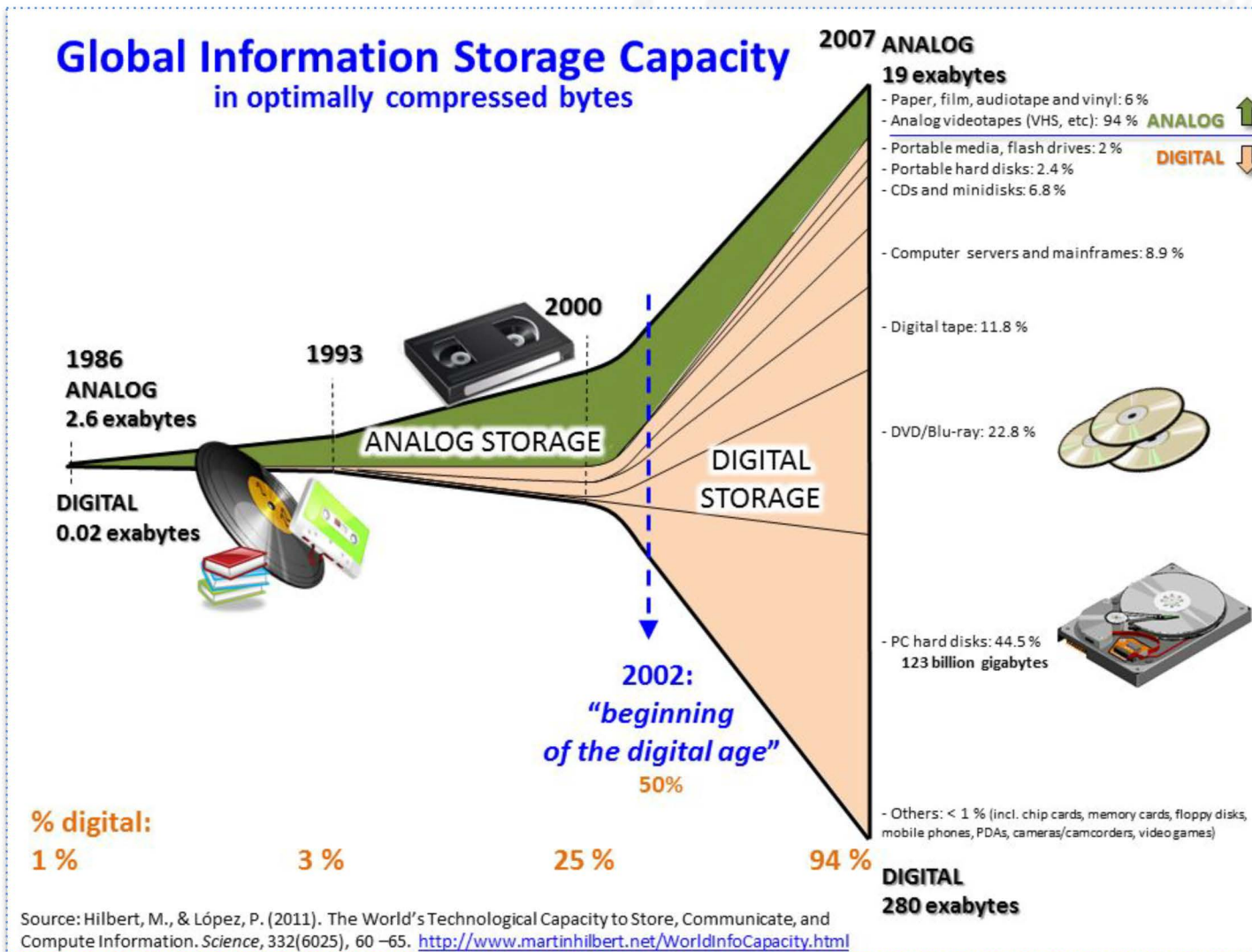


SHARE & PUBLISH



MANAGE, STORE, PRESERVE

A brief history of the 'data deluge'



Storage & Backup ≠ Preservation

Yes, maintaining **backups** of your stored data is crucial! But this 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 generally handled by specialist staff, such as archivists and librarians, using dedicated hard- and software solutions. Researchers need to be aware that some of their data may legally require digital preservation, and ideally participate actively in the process of planning for it from the outset (see: DMP).

Example: Digitisation for digital preservation

'legacy' media	hardware	software	digital files		
			master (preservation)	service (working)	access (access)
<u>Documents</u> : manuscripts; theses; ...	flatbed scanner; feeder scanner; ...	Acrobat Pro;tif	.jpg	.jpg .txt .pdf
<u>Images</u> : 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
<u>Audio</u> : 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
<u>Video</u> : Umatic; Betacam; VHS; MiniDV; ...	Umatic, Betacam and VHS cleaners and recorders; MiniDV, DVCam descks; ...	MediaExpress; FinalCut Studio; Premiere Pro;mov .mpg2	.avi	.mp4





arkivum
Bringing archived data to life



DIGITAL LIBRARY
SERVICES

A future problem: **where is my data?**

I know where it is 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

I don't even know where it is...

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>



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



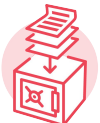
COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



arkivum
Bringing archived data to life



DIGITAL LIBRARY
SERVICES

Digital preservation technologies & processes

File format
normalisation

ESCROW*

Migration
paths

Multiple
copies

Automatic
metadata
capture

Fixity
checking
& virus
scans

Search
platform

Data
under
managem
ent

* **Source code escrow** is the deposit of the [source code](https://en.wikipedia.org/wiki/Source_code_escrow) of [software](https://en.wikipedia.org/wiki/Source_code_escrow) with a third-party [escrow](https://en.wikipedia.org/wiki/Source_code_escrow) agent. Escrow is typically requested by a party licensing software (the licensee), to ensure maintenance of the software instead of [abandonment](https://en.wikipedia.org/wiki/Source_code_escrow) or [orphaning](https://en.wikipedia.org/wiki/Source_code_escrow).⁷ Online. Available: https://en.wikipedia.org/wiki/Source_code_escrow

Adapted from: Arkivum: **Webinar Recording - Making the case for digital preservation**. Available:

<https://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>



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



Closing Remarks & Future Interactions



DS, RDM and OS

Digital Scholarship is the application and integration of digital tools and methods discover, research and teach.

Research Data Management is the organization and documentation of research data (ideally towards making it **Findable**, **Accessible**, **Interoperable** and **Reusable**).

Open Science is a set of *practices* that drives all aspects of research to be more efficient, accountable, collaborative, and of good quality.

Digital Scholarship, within a research project, integrates digital technologies, works within networked environments and subscribes to Open Science. When all the above intersect, DS has the power to transform the research landscape.

Upcoming Workshops

RESEARCH DATA MANAGEMENT TRAINING

ALL SESSIONS @ 10AM IN ULWAZI TRAINING ROOM

Discover how you can become a more **EFFICIENT** researcher in today's digital world. Start managing your **DATA** and your **RESEARCH** process with guidance from the **DLS TEAM**.

RESEARCH DATA MANAGEMENT WITH DMPONLINE



PLAN & DESIGN

The new Student MoU as well as the NRF require students to outline their data plans for their research projects in a Data Management Plan (DMP). This talk/workshop takes you through the reasons for creating a DMP, as well as guiding you through using the DMPonline website.



WEDNESDAY
12 JUN | 14 AUG

DOING DIGITAL SCHOLARSHIP



COLLABORATE & ANALYSE



COLLECT & CAPTURE

Doing research requires interacting with a multitude of digital spaces. This talk outlines digital processes and tools that can increase efficiencies throughout a research project. It looks at collaborative tools for managing, analyzing, mapping and visualizing research data.

WEDNESDAY
15 MAY | 11 SEP | 13 NOV | 11 DEC

SHARING AND PUBLISHING WITH ZIVAHUB



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



UCT's open data repository is rapidly growing. Uploading your research outputs to **ZivaHub** makes them discoverable, citable, shareable and reusable. Learn about open data and **ZivaHub** which allow you to engage with researchers at UCT and the world.

WEDNESDAY
10 JUL | 9 OCT

Research Data Management:
Plan for your Data with DMPonline
12 June @ 10 AM



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD



PLAN & DESIGN



COLLECT & CAPTURE



COLLABORATE & ANALYSE



DISCOVER, REUSE & CITE



SHARE & PUBLISH



MANAGE, STORE, PRESERVE



'RDM at UCT' Slack workspace

Slack: 'Searchable Log of All Conversation and Knowledge'

RDM at UCT v
Niklas Zimmer

Jump to...

All Unreads
All Threads

Starred
rds-wg

Channels

- # academia-edu
- # carpentries
- # data-curation
- # digital-preservation
- # digital-scholarship
- # dls_engage
- # dmp
- # g-suite
- # general
- # gis
- # identifiers
- # linked-data
- # metadata
- # oer
- # openness
- # osf
- # random
- # rdm_newsfeeds
- # rdm_rds
- # software-as-research
- # zenodo
- # zivahub_figshare

#zivahub_figshare

Monday, June 11th

manual – for accessing discovered data; a requirement to openly and richly describe the context within which those data were generated, to enable evaluation of its utility; to explicitly define the conditions under which they may be reused; and to provide clear instructions on how they should be cited when reused. None of these principles necessitate data being “open” or “free”. They do, however, require clarity and transparency around the conditions governing access and reuse. As such, while FAIR data does not need to be open, in order to complete the condition of reusability, FAIR data are required to have a clear, preferably machine readable, license. The transparent but controlled accessibility and services, as opposed to the ambiguous blanket-concept of “open”, allows the participation of a broad range of sectors – in equal partnership with stakeholders in all sectors.

<https://content.iospress.com/articles/information>

Wednesday, July

Niklas Zimmer 10:26

https://figshare.com/articles/Monash_University_study/6396776

Monash University's Content Migration: A case study

Paper posted on 31.05.2018, 15:24 by Andrew Harrison

This is a case study based on Monash University's experience migrating content, including their theses, into their instance of Figshare.

For more information on Monash University's content migration, including the coding required to migrate the content and work done in-house versus commissioned by the university, please reach out to Andrew: andrew.harrison@monash.edu.

References

<https://monash.figshare.com/theses>

figshare

Monash University's Content Migration: A case study

This is a case study based on Monash University's experience migrating content, including their theses, into their instance of Figshare. For more information on Monash University's content migration, including the coding required to migrate the content and work done in-house versus commissioned by the university, please reach out to Andrew: andrew.harrison@monash.edu.

Related channels

72 Members

- Andre Le Roux
- Andrew Deacon

[RDM at UCT \(Slack\)](#)

[UCT DMPonline](#)

[OneDrive / Google Drive etc.](#)

[UCT Open Science Framework \(OSF\)](#)

[Digital preservation](#)

[ZivaHub | Open Data UCT](#)



Declaration On Research Assessment

Improving how research is assessed

sfdora.org

@DORAssessment

now also
including
UCT!

Signed by >500 organizations and >12,500 individuals

Supporting organizations





Thank you!



dls@uct.ac.za



[@DigitalUct](https://twitter.com/DigitalUct)



rdm-at-uct.slack.com



DIGITAL LIBRARY
SERVICES

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