Using Figshare
a quick demonstration for the ilifu user group

Friday, 29th March 2019, 13:45 - 14:00
CPUT Bellville Library Conference Centre

Niklas Zimmer
Manager
Digital Library Services

https://doi.org/10.25375/uct.6264638
I hold an MA(FA) and a BA(Hons) from University of Cape Town, as well as a BA in education from University of Cologne. My research MA on South African Jazz photography drew critical reflections on Fine Arts, Social history and Music. Before my employment at UCT Libraries in 2015 as manager of the Digital Library Services department, I worked as digitisation manager at the Centre for Popular Memory (UCT), as an art teacher, heading the Visual Arts Department at the German International School Cape Town, and as an archivist at the Art and Exhibition Hall of the Federal Republic of Germany. I have maintained a richly diverse freelance career, including lecturing in theory and discourse of art and critical studies at tertiary institutions in Cape Town and Stellenbosch, giving workshops in video, sound and photography, as well as writing reviews, exhibiting my photographic art. I am a research associate at Africa Open Institute for Music, Research and Innovation at Stellenbosch University.
FREYA is a 3-year project funded by the European Commission under the Horizon 2020 programme. The project aims to build the infrastructure for persistent identifiers as a core component of open science, in the EU and globally. FREYA will improve discovery, navigation, retrieval, and access of research resources. New provenance services will enable researchers to better evaluate data and make the scientific record more complete, reliable, and traceable. By engaging with the global community through the Research Data Alliance and other research infrastructures, we work together to realise the vision of fully and effectively accessible data.
1. DMPs | Resourcing and planning for data publishing
2. 5-star Open data | FAIR data
3. How to upload and publish data on Figshare
4. Data availability statement | DOI
5. Groups in Figshare
6. Tiered instances of Figshare
7. DLS resources on RDM at UCT Slack Workspace
DMPs | Resourcing and planning for data publishing

Plan

Process

Analyse

Preserve

Publish

Reuse

Proposal

MoU

DMP (proposal)

“I will do XYZ to manage my data”

DMP (active)

“I am doing XYZ to manage my data”

DMP (final)

review

“ I have done XYZ to manage my data”

DMP ONLINE

ZivaHub

UNIVERSITY OF CAPE TOWN
# The Support-Your-Data RDM rubric

<table>
<thead>
<tr>
<th>1. Planning your project</th>
<th>Ad Hoc</th>
<th>One-Time</th>
<th>Active and Informative</th>
<th>Optimized for Re-Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>When it comes to my data, I have a &quot;way of doing things&quot; but no standard or documented plans.</td>
<td>I create some formal plans about how I will manage my data at the start of a project, but I generally don't refer back to them.</td>
<td>I develop detailed plans about how I will manage my data that I actively revisit and revise over the course of a project.</td>
<td>1. I have created plans for managing my data that are designed to streamline its future use by myself or others.</td>
<td></td>
</tr>
</tbody>
</table>

| 2. Organizing your data | I don't follow a consistent approach for keeping my data organized, so it often takes time to find things. | I have an approach for organizing my data, but I only put it into action after my project is complete. | I have an approach for organizing my data that I implement prospectively, but it not necessarily standardized. | 2. I organize my data so that others can navigate, understand, and use it without me being present. |

| 3. Saving and backing up your data | I decide what data is important while I am working on it and typically save it in a single location. | I know what data needs to be saved and I back it up after I'm done working on it to reduce the risk of loss. | I have a system for regularly saving important data while I am working on it. I have multiple backups. | 3. I save my data in a manner and location designed maximize opportunities for re-use by myself and others. |

| 4. Getting your data ready for analysis | I don't have a standardized or well documented process for preparing my data for analysis. | I have thought about how I will need to prepare my data, but I handle each case in a different manner. | My process for preparing data is standardized and well documented. | 4. I prepare my data in such a way as to facilitate use by both myself and others in the future. |

| 5. Analyzing your data and handling the outputs | I often have to redo my analyses or examine their products to determine what procedures or parameters were applied. | After I finish my analysis, I document the specific parameters, procedures, and protocols applied. | I regularly document the specifics of both my analysis workflow and decision making process while I am analyzing my data. | 5. I have ensured that the specifics of my analysis workflow and decision making process can be understood and put into action by others. |

| 6. Sharing and publishing your data | I share the results of my research, but generally I do not share the underlying data. | I share my data only when I'm required to do so or in response to direct requests from other researchers. | I regularly share the data that underlies my results and conclusions in a form that enables use by others. | 6. Because of my excellent data management practices, I am able to efficiently share my data whenever I need to with whomever I need to. |

---


NB: these example workflow stages are probably pretty wrong … :)

---

[Digital Library Services](https://digital.library.universitas.edu)
5-Star Open data | FAIR data
5 ★ Open Data

★
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/

Friday, 29th March 2019
FAIR data

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

Findable

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

Interoperable

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

Accessible

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

Reusable
How to delete my account?

How to choose the most appropriate licence
A guide to licensing your data, for institutional users only.

How to edit or delete my data
Find out how to make changes to your data. Please note that only private data can be deleted.

How to upload confidential files, linked files, embargoed files, and metadata records only
There are many ways to upload your data and/or metadata.

How to sync ORCiD and DataCite for figshare

How to connect figshare with your GitHub account

How to use Collections

How to use projects and collaborate on Figshare

How to get a Personal Token

How to use the OAI-PMH
Navigate to your institutions’ Figshare portal and select ‘Log in’ at top right and enter your SSO (Single Sign-On) details. The process may vary depending on the respective institutional setup, but most Figshare instances in South Africa will likely be using SAFIRE for authentication.

Source: https://knowledge.figshare.com/
Select the **Create a new item** button from the top left (or just drag and drop your file into the browser). If you are uploading a big file size then you can use our desktop uploader, see: [how to use the desktop uploader](https://knowledge.figshare.com/).
Insert a Title

For information on how best to complete the fields for maximum exposure of your data, see: Figshare how-to guide. Tips on how to fill out each of the fields can also be found on the right side of the screen.

Source: https://knowledge.figshare.com/
Add Authors

You can rearrange the order in which the authors appear and remove yourself as an author if you are uploading on behalf of someone. You can also search by entering the email address of the author you want to add.
Select a **Category**

These are taken from the Australian Fields of Research classification system. You can choose more than one and either select from the drop-down menu or search for your subject area. It is suggested to find the closest fit, and then get more specific in the Keywords section, as required.

Source: [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
Select the File Type

Select the File type from the drop-down menu. Figshare accepts over 1,000 file extensions, and many are parsed right in the browser. If there’s something that isn’t supported yet, they will likely build it for you.

Source: https://knowledge.figshare.com/
Add Keyword(s)

These should be more specific than the category and help others find your research. Add as many keywords as you want - just hit enter after each keyword. figshare remembers ones you have entered previously, and will also suggest ones. **NB: please include ‘ilifu’ here.**

Source: [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
Write a **Description**

‘The abstract’ - include any relevant information that pertains to your research, e.g. information about the methodology, approval for data collection, or legal or ethical requirements. Be as descriptive as you can!
Add Funding info

When typing in your funder information, Figshare will search the Dimensions database for your funder and hyperlink to the Dimensions page bridging the connection between your research outputs and the funder.
Select a License

You can choose from a number of licenses based on your reuse requirements. For more information see: How to choose the most appropriate licence.

Source: https://knowledge.figshare.com/
1. You can copy & paste URLs into the **References** box (add more fields by hitting Enter).

2. You can link to data that are hosted elsewhere by selecting ‘**Link file**.’

Source: [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
Tick Publish

Make sure everything you have submitted is accurate - once it’s published, it’s **permanently available**. You can go back and edit items after you’ve made them publicly available. Some changes may trigger a new version. See [here](https://knowledge.figshare.com/) to find out which amends will generate a new version of your item.
All of your items will first be reviewed by curator before they are approved (i.e. published). Someone may contact you, generally with suggested edits towards improving the metadata. In some cases, they may also take down files that have been published in error.
Those are the basics, and now for a few more features ...
Confidential data files are securely stored, but are inaccessible to the user. Just the metadata, i.e. the item record, is open. This is an option for ethically or commercially sensitive data.

Source: https://knowledge.figshare.com/
**Embargoing data** serves to allow the author(s) a period of sole use of the data. After you click ‘save’ & ‘publish’ - and approval by the curator -, users are presented with a countdown timer.

Source: [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
Publishing a **Metadata-only record**

In this case no files are uploaded, but a description of the study is available, including information about how to contact the authors. This is a good option for getting a DOI for - and linking to - data that are stored / hosted somewhere else.

Source: [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
Creating a Project

Projects are **collaborative spaces used for ongoing work**. You can upload data that is in progress and have users make comments. **Projects are secure spaces that can be used for sensitive data.** You can also **collaborate with people outside your institution** by inviting them to your project.

Source: [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
## Individual vs. Group Projects

If a project is for individual use, select the **individual** option under Allocate storage. To use group storage, select the **group** option and find the relevant group. Add users to the project by searching on the right hand side (collaborator or viewer). Collaborators can comment on the project and the data within it and upload items, while viewers can only view the data. NB: Items must be downloaded, edited, and re-uploaded to the Project.

<table>
<thead>
<tr>
<th>Individual Projects</th>
<th>Group Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Everyone uses their own quota and account storage.</td>
<td>Submitters’ quota will not be used, storage allocation comes directly from the project.</td>
</tr>
<tr>
<td>People take their work with them if they leave the project.</td>
<td>All work is stored on institutional storage and remains within the project space if people leave.</td>
</tr>
<tr>
<td>Items are created using the metadata schema of the submitter.</td>
<td>Contributors must adopt the metadata schema of the project owner.</td>
</tr>
<tr>
<td>Items appear in the subgroup of the uploader.</td>
<td>Items appear in the subgroup of the project owner.</td>
</tr>
</tbody>
</table>

Source: https://knowledge.figshare.com/
Creating a **Collection**

Collections are ways of collating data that bring it together under a theme. They can be either private or public and can be assigned a DOI. One key difference the Projects feature is that this offers a way of gathering and re-using data published on Figshare which you did **not** create yourself.

**Source:** [https://knowledge.figshare.com/](https://knowledge.figshare.com/)
+Add items to a Collection

Once you’ve created a Collection, select Add public items or My data. If you select from public items, you can search and select the items you wish to add to your collection. The same process applies if you choose from your data. If you find a public item you wish to add to a Collection, simply select +Collect and choose the Collection to add it to.

Source: https://knowledge.figshare.com/
Once you’ve collected your data, you can make your Collection public by selecting the Manage gear wheel on the right side of the screen and selecting Publish collection. Once you’ve published a collection, it’s permanent - the same as publishing an item.

Source: https://knowledge.figshare.com/
Data availability statement | DOI
The ilifu Research Data Management and Open Science Guidelines (V.0.5 June 2018) require users to share their data openly by default (with various exceptions provided for). This also requires users to provide a link to where the data underpinning their traditional, written research outputs can be found. Provided are a few examples of wording data availability statements based on whether:

1. you collected or generated your own data
2. you re-used publicly-available data
3. your data requires an embargo period, or
4. (if applicable), why your data is not publicly available.

Data availability statements should be included in the abstract or executive summary, as well as in the methods/methodology section. When data is uploaded to Figshare, a doi is automatically generated on saving the new item. This doi can be copied & pasted into the data availability statement.
The data supporting this paper can be accessed on the University of Cape Town’s Institutional Data Repository (ZivaHub) under the following doi: ............................
The data that support the findings of this study are openly available in {repository name, e.g. ‘ZivaHub’} under the following {identifier(s), e.g. DOI(s)}: ‘…’

Example from the Department of Biological Sciences:

The data that support the findings of this study are openly available in ZivaHub with the identifier 10.25375/uct.6561776
The data that support the findings of this study are available in [repository name] at [URL/DOI], reference number [reference number]. These data were derived from the following resources available in the public domain: [list resources and URLs]

Example from the Faculty of Commerce:

The data that support the findings of this study are openly available in DataFirst with the identifier

https://datafirst.uct.ac.za/dataportal/index.php/catalog/570
Data availability statement

(3) You collected your own data, but require an **embargo period**:

The data that support the findings will be available in [repository name] at [URL / DOI link] following a [6 month] embargo from the date of publication to allow for commercialisation of research findings.

Example from Faculty of Science:

The data that support the findings will be available in ZivaHub at 10.23456/uct.examplelink following a 12 month embargo from the date of publication to allow for commercialisation of research findings.
Data availability statement

(4) Ethical concerns *prevent* data sharing:

Note: even if your data cannot be ethically shared, the metadata describing the study should still be made available as it builds your academic profile and may lead to future collaboration. UCT’s data repository **ZivaHub** ([https://uct.figshare.com/](https://uct.figshare.com/)) allows the creation of *metadata-only records* that can be used to describe a study without openly sharing data objects.

**Example:**

*Due to ethical concerns, supporting data cannot be made openly available.*

*Further information about the data and conditions for access are available at the University of Bath data archive: [https://doi.org/10.15125/12345](https://doi.org/10.15125/12345)*

If none of the above adequately describe your data sharing requirements, the University of Bath has comprehensive guidelines ([http://www.bath.ac.uk/research/data/archiving-data/writing-a-data-access-statement/](http://www.bath.ac.uk/research/data/archiving-data/writing-a-data-access-statement/)) on how to write different kinds of data accessibility statements.
The (institution-specific) doi on Figshare

The dois created and managed via Figshare are minted by DataCite. A doi is an alphanumeric string beginning with ‘10’, i.e 10.25375/uct.7143581.v1. To make this resolve to a digital object, it needs to be prefixed with ‘https://doi.org/’.

Example: [https://doi.org/10.25375/uct.7143581.v1](https://doi.org/10.25375/uct.7143581.v1), further explained:

<table>
<thead>
<tr>
<th>https://</th>
<th>doi.org/</th>
<th>10.25375/</th>
<th>uct.</th>
<th>7143581.</th>
<th>v1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A secure extension of HTTP.</td>
<td>International DOI foundation, who ensure that DOIs stay true.</td>
<td>‘10’ is the standard doi directory code; ‘25375’ is the registrant code for ZivaHub.</td>
<td>Indicates that the DOI originates from a UCT researcher or research project.</td>
<td>Administrative subdivision code.</td>
<td>Version number.</td>
</tr>
</tbody>
</table>
Groups in Figshare
2. Discover research from Monash University Theses

3. Faculty of Science Theses

Source: https://monash.figshare.com/sciencetheses

level 2 group
Discover research from Faculty of Science Theses

3.

level 3 group
Tiered instances of Figshare

**Tier 1:** national - DIRISA storage? curation?

**Tier 2:** regional - ilifu instance? group?

**Tier 3:** institutional - ilifu sub-groups? projects?

**Tier 4:** individual - ilifu affiliates? tags?
National Non-Traditional Research Output Portal

Region or subject specific 'Data' Portals

Portals for each grant code

Autopopulated

Autopopulated

e.g. data.tnderx.com

e.g. genomics.tnderx.com

Normalized and mapping to Funded Grant Code & Institution using GRID, UberResearch & FundRef

Funded Researcher Uploads

Academic Publisher Datasets from e.g. SpringerNature, ACS, Wiley etc

Migration or ingest of data from other funded Repositories, or Institutional Data Repositories

Migration or ingest of data from popular repositories e.g. Figshare, Dryad, Zenodo

DLS resources on RDM at UCT Slack Workspace
RDM at UCT (Slack)

UCT DMPonline

OneDrive / Google Drive etc.

UCT Open Science Framework (OSF)

Digital preservation (DLS)

ZivaHub | Open Data UCT
Thank you