



# Writing a good Data Management Plan (DMP): A workshop

Nicholas Syrotiuk  
Research Data Manager  
Durham University

15 November 2023



# Session plan

Introduction to  
RDM + Task 1

20:00 [1]

Data manage-  
ment planning  
+ Task 2

30:00 [2]

Demo of  
DMPonline tool

10:00 [3]

Try writing  
a DMP

45:00 [4]

Part one:

# INTRODUCTION TO RDM

What do we mean by

# RDM?

RESEARCH

DATA

MANAGEMENT



RDM helps to preserve, protect and proliferate the data behind scientific (research) discoveries and claims

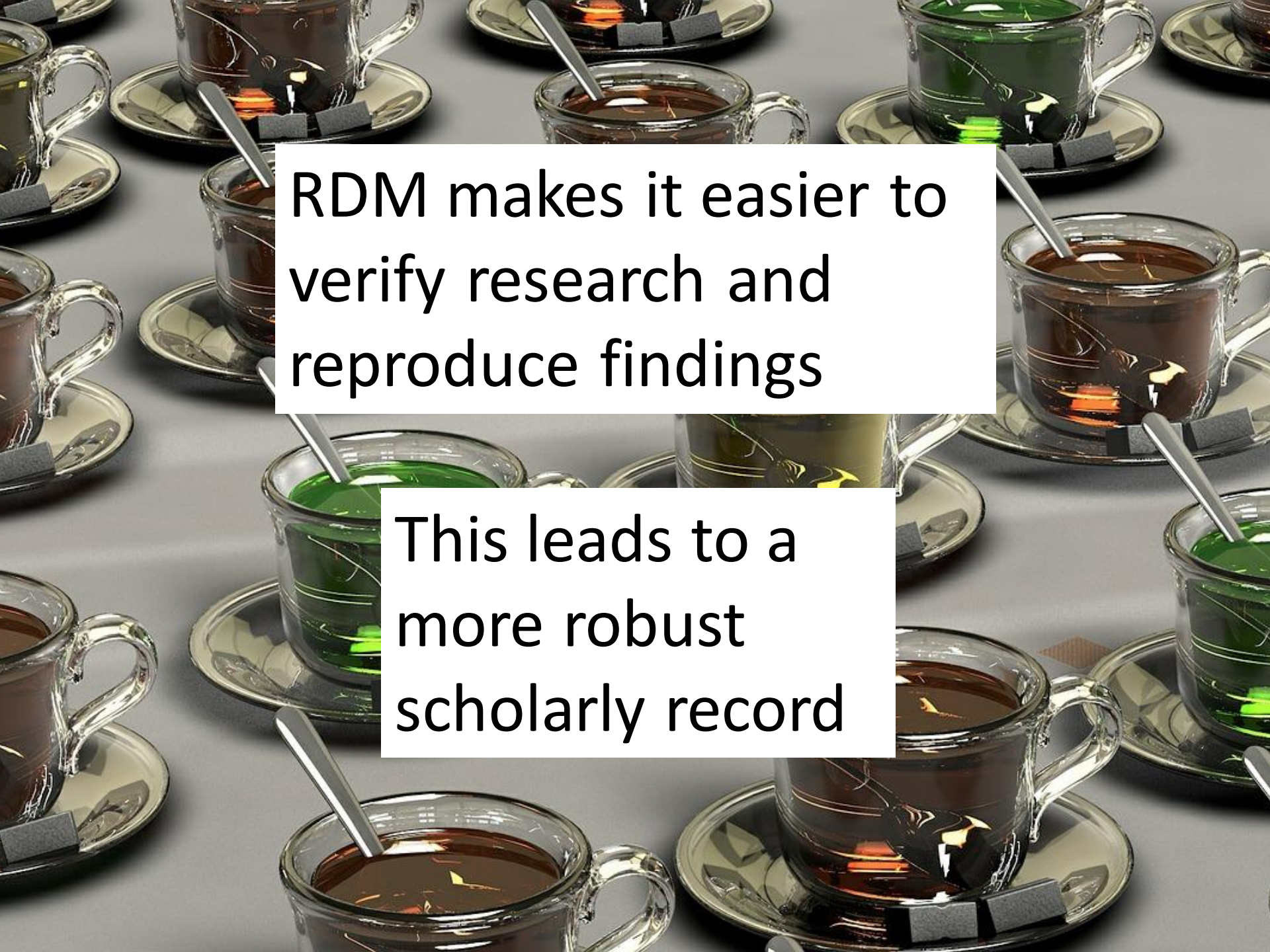
The aim is high quality RDM



RDM leads to  
increased  
transparency  
of the research  
process







RDM makes it easier to  
verify research and  
reproduce findings

This leads to a  
more robust  
scholarly record

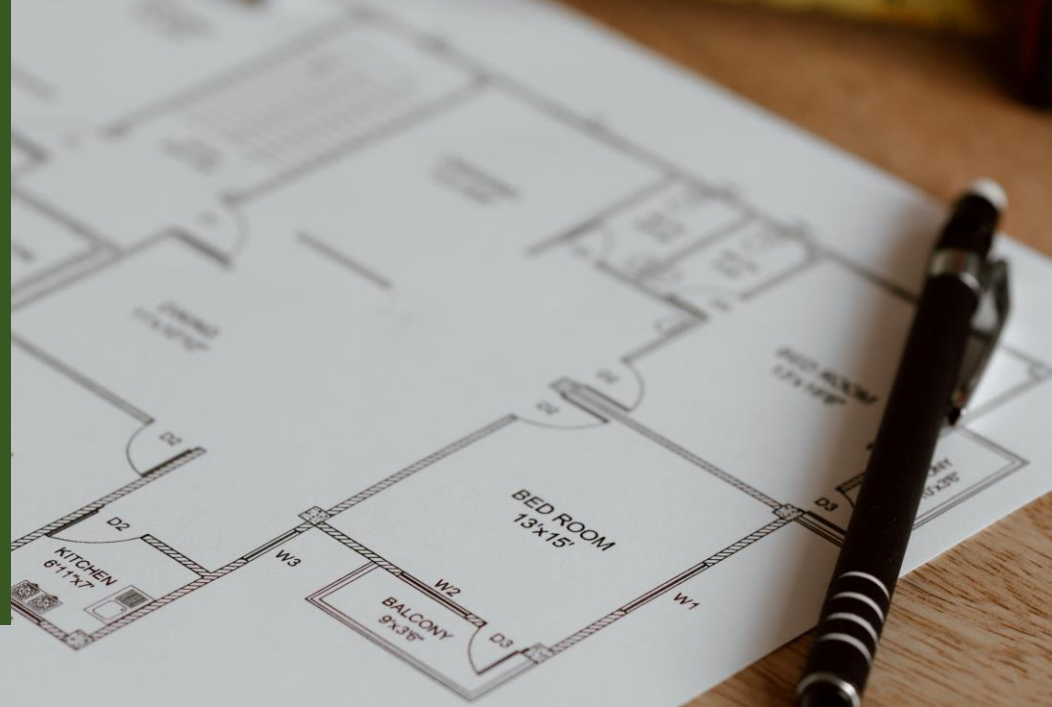
RDM and sharing  
(anonymised) research  
data can lead to making  
more progress as  
a research community  
collectively




What are  
core RDM  
activities?



1. Planning  
and describing  
data-related  
work before it  
takes place



A close-up, low-angle shot of a thick stack of old, yellowed papers. The pages are slightly curved and layered, creating a sense of depth and texture. The lighting is warm, highlighting the golden-brown tones of the paper. A semi-transparent blue rectangular box is overlaid on the right side of the image, containing white text.

## 2. Documenting your data, your data processing and your workflows



### 3. Choosing open or "standardised" file formats where possible

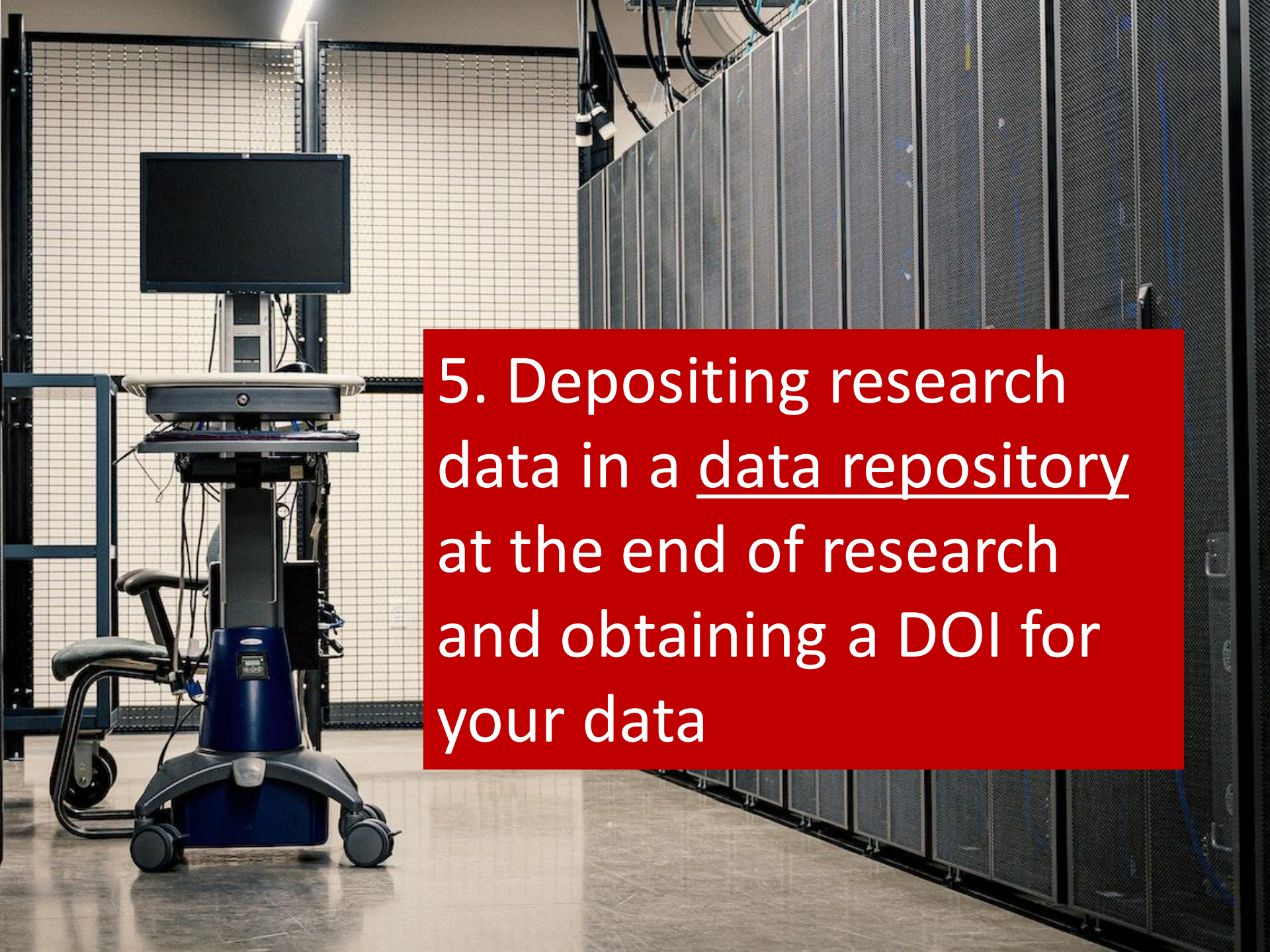






## 4. Storing research data safely during a project



A photograph of a server room. On the left, there is a workstation on a blue mobile cart with a monitor, keyboard, and mouse. The cart is on wheels. To the right, there are several rows of server racks with mesh doors. The floor is a light-colored, polished surface. The background shows a grid-like structure, possibly a wall or ceiling panel.

5. Depositing research data in a data repository at the end of research and obtaining a DOI for your data



6. Linking publications to the datasets that underpin them and increasingly code/scripts too



# Why bother with data management?

An aerial photograph of a large, multi-story research facility. A central building is engulfed in flames, with thick black smoke billowing upwards and spreading across the sky. The surrounding area includes other large buildings, parking lots, and green spaces. The overall scene depicts a major disaster at a research center.

2005/10/30:

Fire destroys Southampton research centre

2019/01/28:

Reward for Royal Oak wreck  
data (laptop and backup  
discs) stolen from flat



# 150,000 arrest records wiped in tech blunder

Offenders may go free after software bug deletes fingerprint and DNA files on police computer

Fiona Hamilton, Crime and Security Editor

Friday January 15 2021, 12.01am,  
The Times

UK politics

Politics



The error may allow offenders to go free because biometric evidence left at crime scenes will not be flagged up  
GETTY IMAGES

Source: The Times online (15 Jan 2021)



Lost five  
years of  
research  
data

# CASH REWARD

for returning my lost backpack



*303 Adventure.com*

- Black [AK] Burton Rucksack
- Lost on Friday 15. July at 8 pm in the Panton Arms pub 43, Panton St. Cambridge
- Containing a laptop (white MacBook), a black external hard drive and scientific research documents

The external hard drive is VERY important to me as it contains 5 years of research data which are crucial for my PhD thesis!!!

If you found it, I would be extremely grateful if you could return it to the Panton Arms or contact me on: [REDACTED]

Thank you!!



DOES ANYONE HAVE A  
HORROR STORY ABOUT  
DATA LOSS?

# Types of research data

Survey data

Recordings

Text documents

Tabular data

Photographs

Image data

Big, new, novel or voluminous data

Scientific measurements

MRI data in BIDS format

NVivo data

International macrodata

Census data

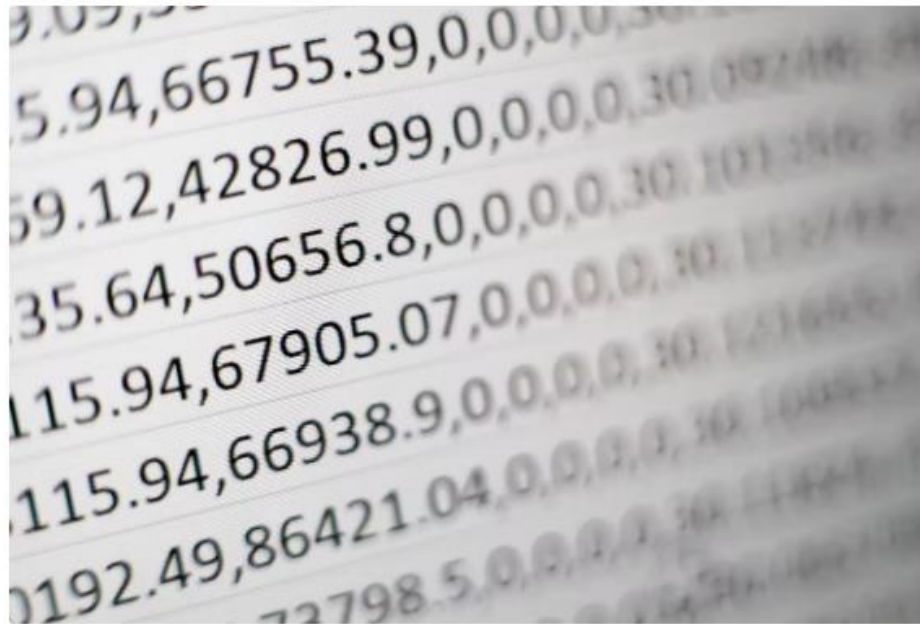
Code / scripts

Project web site

Task one:

WHAT TYPE(S) OF  
RESEARCH DATA DO  
YOU CREATE?

# Mentimeter



Name one or two types of research data you collect or create.

25

25

Submit

Name one or two types of research data you collect or create.

39 responses





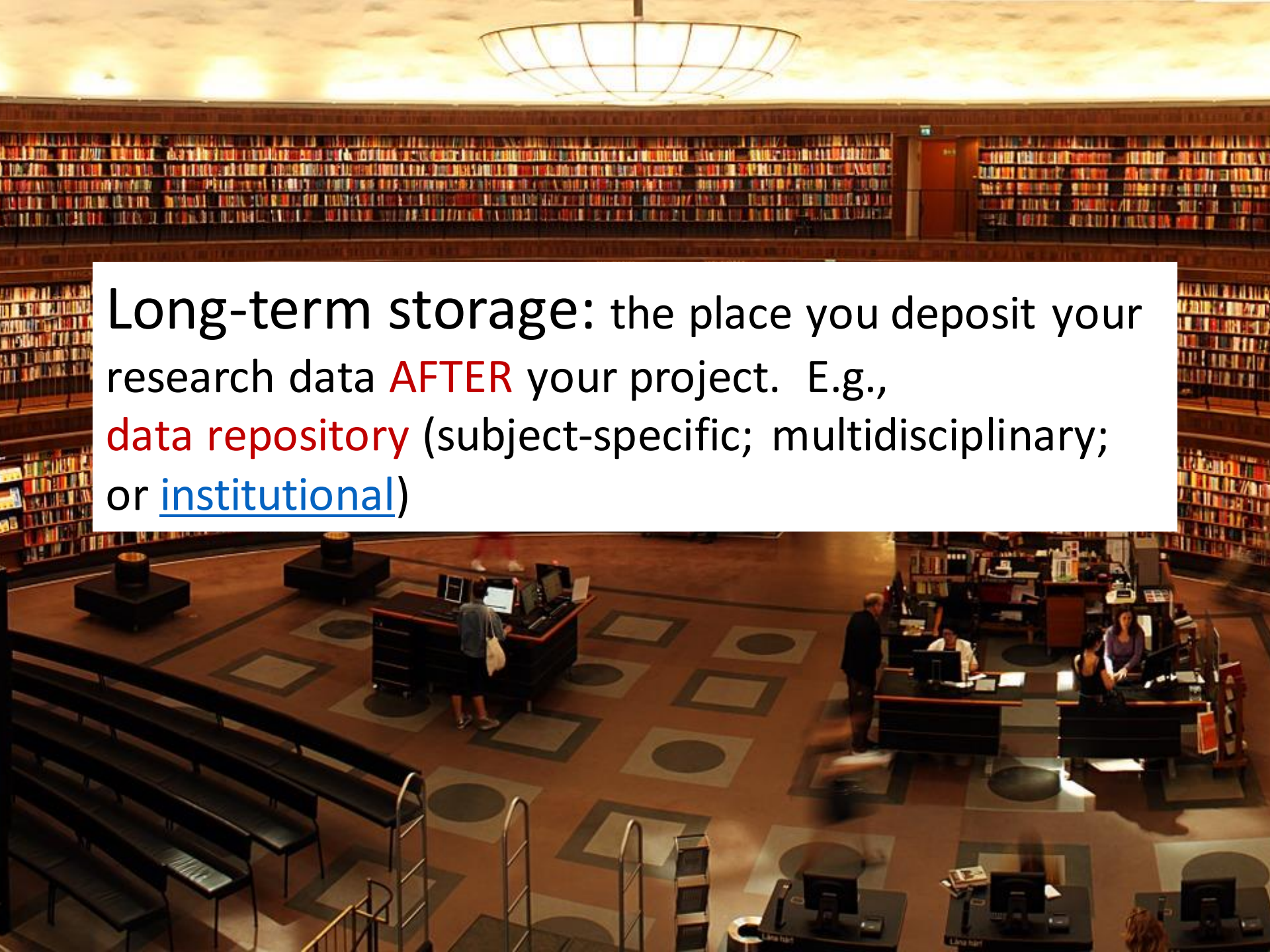
A crowd of people is gathered in a dark room, looking towards a glowing blue sign that reads "TERMINOLOGY". The sign is illuminated with a bright blue light, and the word is written in a white, blocky font. The room is dimly lit, with red vertical light strips on the walls and a blue glow from the sign. The crowd is silhouetted against the light, and the overall atmosphere is mysterious and futuristic.

TERMINOLOGY

**Short-term storage:** the place you keep your research data DURING your project. E.g., OneDrive for Business or Shared Research Storage (SRS)







Long-term storage: the place you deposit your research data **AFTER** your project. E.g., **data repository** (subject-specific; multidisciplinary; or institutional)



...gine a revenue of hundred million dollars

### Global Finance

Resource	Balance Value(\$)
<b>Loans Portfolio</b>	2,568,790
- Auto Loans	876,510
- Personal Loans	889,870
- Credit Cards	120,000
- Other Loans	422,410
- Other Assets	79,900
<b>Asset Liabilities</b>	890,144
- Cash	178,990
- Accounts Payable	98,420
- Other Liabilities	244,220
- Other Assets	78,514
<b>Equity/Retained Earnings</b>	144,579
- Retained Earnings	227,489
- Other Assets	143,985
- Other Liabilities	42,900
<b>Shareholders' Equity</b>	48,771
- Common Stock	162,790
- Retained Earnings	279,380
<b>Total Resources</b>	4,567,264

### Stock Market Strategy

Marketing strategy's goal is to increase sales and achieve advantage over other competitors. It includes short term and long term activities of marketing that can be analyzed through analysis of company's situation and contribute to its objectives. The strategy section focuses on how you gain sales by acquiring and keeping customers. A marketing strategy involves making good messages with the right level of marketing approaches to obtain from your customers of your sales and marketing activities. Putting your strategy into action shows your marketing plan should work. Marketing budgets will be set, of the same time, it will show you how you're going to work with your targets, etc.



# THE DAILY

## RESEARCH DATA PUBLISHED!

### NEW RESEARCH DATA PUBLISHED TODAY

#### US RATE RISE CHANCES REVEALED AS JOBS GROWTH SLOWS

#### Collections



#### Recently Uploaded

**Finding gain, plans abound...**

...the world is messed up

**Data sharing:** often means publishing anonymised research data in a data repository\*\* AFTER your research. Other meanings also.

\*\* A repository typically has three levels of access: open access; access for registered users; and restricted access (i.e., access may be granted upon request).

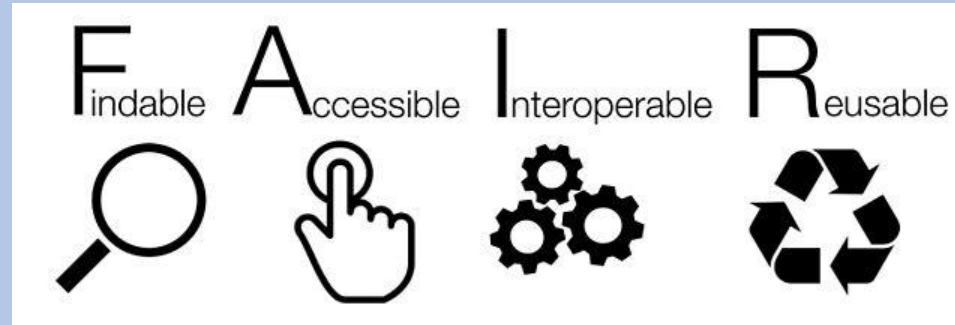




**Data sharing agreement: a legal agreement between two organisations for sharing personal data. Read *Data sharing code of practice* from ICO. Seek legal advice.**



# F.A.I.R. data is ...



The aim is to create **F.A.I.R. data**  
as opposed to unFAIR data

Excellent talk by Henry Rzepa (Imperial College) to  
Durham researchers. DOI: <http://doi.org/cppz>

Part two:

**DATA**

**MANAGEMENT**

**PLANNING**




UKRI COMMON PRINCIPLES  
ON DATA POLICY

UKRI provides guidance on best practice in the management of research data




Plan to keep  
your sensitive  
data safe



A blue recycling bin with a white recycling symbol on the side. The bin is empty and has a white text box on the right side.

Plan to maximise the  
re-use potential of  
your data



A high-angle, wide shot of a large, multi-level library. The library features circular bookshelves on multiple levels, creating a tiered effect. The ground floor has a central study area with several computer workstations and people working. The upper levels are filled with bookshelves. The lighting is warm and ambient, with a large, circular, grid-patterned light fixture hanging from the ceiling. The overall atmosphere is quiet and scholarly.

Plan to preserve  
your data for the  
long term



# Structure of a DMP

Description of project

Detailed description of data

Data documentation

Sensitivity of data

Short-term storage of data

Long-term storage of data

# Describe research data in a table

Data created or collected	Data type	Data format	Volume or duration	Planned storage or access
Raw ethnographic field notes	Notebooks	Paper	n/a	No shared access
Photographs	Digital images	JPEG	100 photographs * 4 Mb each = ca. 400 Mb	<a href="#">OneDrive for Business</a>
Interviews	Sound recordings	MP3	20 interviews, ca. 30 minutes each	Destroyed after transcription
Transcriptions	Document	Word	20 documents, 5 Mb each. Total: approx. 100 Mb	Anonymised transcripts to be published in the <a href="#">Durham research data repository</a>
Magma flows	X-ray images	Tiff	50 Tb storage per year for five years = 250 Tb total storage	<a href="#">Shared Research Storage</a>



## Example row in table

Data created	Data type	Data format	Volume or duration	Planned access
Magma flows	X-ray images	Tiff format	50 Tb storage per year for five years = 250 Tb total storage	<a href="#">Shared Research Storage</a>

Task two:

**DESCRIBE YOUR  
RESEARCH DATA IN A  
TABLE**





TEN RULES  
FOR A GOOD  
DMP



1. Determine  
**FUNDER**  
requirements





## 2. IDENTIFY the data to be collected

Anticipated types of data

Anticipated file formats: open preferred

Anticipated volume of data in bytes

Planned storage and access





3. State how data will be

**ORGANISED**

Spreadsheet

Database

Geographic information system

Nvivo

Qualtrics



Structured metadata e.g., DDI

Comments in code / scripts

README file

4. Explain how data and  
data processing will be  
**DOCUMENTED**



QA during data collection phase

QA during data entry phase

QA during data analysis phase

*High  
Quality*

5. Describe how data

**QUALITY**

will be assured





6. Provide a solid  
**STRATEGY** for  
short-term data  
storage and long-  
term data storage

## 7. Define the project's data

# POLICIES

Data (and software) licencing

Anticipated plans for data sharing

Management of personal data



Deposit data in a research data repository (perhaps with access restrictions)

Submit data to a journal

Publish a data paper

8. Describe how data will be  
**DISSEMINATED**



# 9. Assign ROLES and RESPONSIBILITIES

Guidance on [authorship & acknowledgements](#) in Library guides:

CRediT (Contributor Roles Taxonomy)

Data collector

Quality control

Data analyst

Data visualisation

Data deposit

DMP editor

10. Prepare a realistic

# BUDGET

Software

Hardware

Short-term storage

Tiers: [Silver](#) | [Gold](#)

Long-term storage



Part three:

**DEMONSTRATION  
OF DMPonline**

# Live Demonstration

A large crowd of people is silhouetted against a bright, blue-tinted background, likely a stage or arena. Many individuals have their hands raised in the air, suggesting an active participation or a high-energy event. The lighting is dramatic, with strong highlights and deep shadows, creating a sense of movement and excitement.



Part four: **Try writing a**

**DATA**

**MANAGEMENT**

**PLAN**



## Library Research Support: Open Research: Writing a good Data Management Plan

This guide is intended to provide advice and support on open access research, including guidance around Durham Research Online (DRO), open access publishing, research data management and related topics.

- Home
  - What is Open Access? ▾
  - University & Funder Policies ▾
  - Durham Research Online (DRO) ▾
  - Publisher Deals
  - REF OA ▾
  - PGRs & eTheses
- Research Data Management ▾

### Data management planning

A Data Management Plan (DMP) is a short document which explains to your funder and to your collaborators how you intend to manage your research data during and after your grant funding period. Writing a DMP should not be seen as an administrative burden but rather an opportunity to convince your funder that you have a strategy for managing research data responsibly. What research data will you create? How will you guarantee your research data is stored safely? How will you protect personal data? How will you store large volumes of research data? Where might you publish the research data which supports your findings? You will propose solutions to some of these questions in your DMP.

UKRI has written [Guidance on best practice in the management of research data](#). This is essential reading for all researchers even if your funder is not UKRI. At the end of the UKRI guidance, there is a DMP template which describes what you need to include in your DMP. Please read this valuable guidance.

After reading the guidance, you should write your DMP using the [DMPonline](#) tool. The tool includes [published DMPs](#) but you will need to filter the list by funder because there are too many DMPs to scan. You can also read a sample of *DMPs written by other Durham researchers*; these appear under 'Durham University plans' on your dashboard in DMPonline.

- Please follow these steps to write your plan:
1. Browse to [DMPonline](#)
  2. Create a DMPonline account using your Durham e-mail address
  3. Login with your Durham institutional account credentials. This should link your two accounts.

### Basic structure of a DMP

- A typical DMP has the following structure:
- Describe the research data you will create or collect. Sometimes you can do this in a table (see Table 1 below).
  - Propose a short-term storage solution for your research data **during** your funding period.
  - Explain how you will protect your research data. A separate guide covers [Managing research data during a project](#).
  - Propose a long-term storage solution for your supporting research data (and code) **after** your funding period. Typically, this means [publishing the research data](#) which supports your findings in a data (or code) repository of some kind. A separate guide covers this topic.

European funding bodies will typically ask you to describe your research data and to answer questions about the [F.A.I.R. data principles](#).

Table 1. Different ways of describing research data

Data created or collected	Data type	Data format	Volume or Duration	Planned storage and access
Raw ethnographic field notes	Notebooks	Paper	n/a	No shared access
Photographs	Digital	JPEG	100 photographs * 4 Mb	<a href="#">OneDrive for Business</a>

# Guidance on writing a good DMP

<https://bit.ly/3ykxwrU>

Plan to make data work for you

D  
m  
re

# dmponline.dcc.ac.uk



Sign in

Create account

Forgot password?

Remember email

Sign in

- or -

Sign in with your institutional credentials

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).



59,972 Users



314 Organisations



65,212 Plans



89 Countries





Thank you

Nicholas Syrotiuk

 @DurhamRdm

## References:

Wilkinson, M. et al. (2016): The F.A.I.R. Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018. DOI: <http://doi.org/10.1038/sdata.2016.18>

UK Research and Innovation (2018): "Guidance on best practice in the management of research data." UKRI web site.

CESSDA Training Team (2017 - 2020). CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC. DOI: <http://doi.org/10.5281/zenodo.3820472>

Wilson G. et al. (2017): Good enough practices in scientific computing. *PLoS Comput Biol* 13(6): e1005510. DOI: <http://doi.org/10.1371/journal.pcbi.1005510>

Rzepa, Henry (2018): F.A.I.R. data as a first class citizen in scientific publishing. Imperial College. DOI: <http://doi.org/cppz>