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

Nicholas Syrotiuk
Research Data Manager
Durham University

07 November 2024

DOI: http://doi.org/10.15128/r2fb494855m

Session plan

RDM concepts and terms

15:00 [1]

Data manage-ment planning + Task 1

30:00 [2]

Demo of DMPonline tool

15:00 [3]

Task 2: Try writing a DMP

30:00 [4]

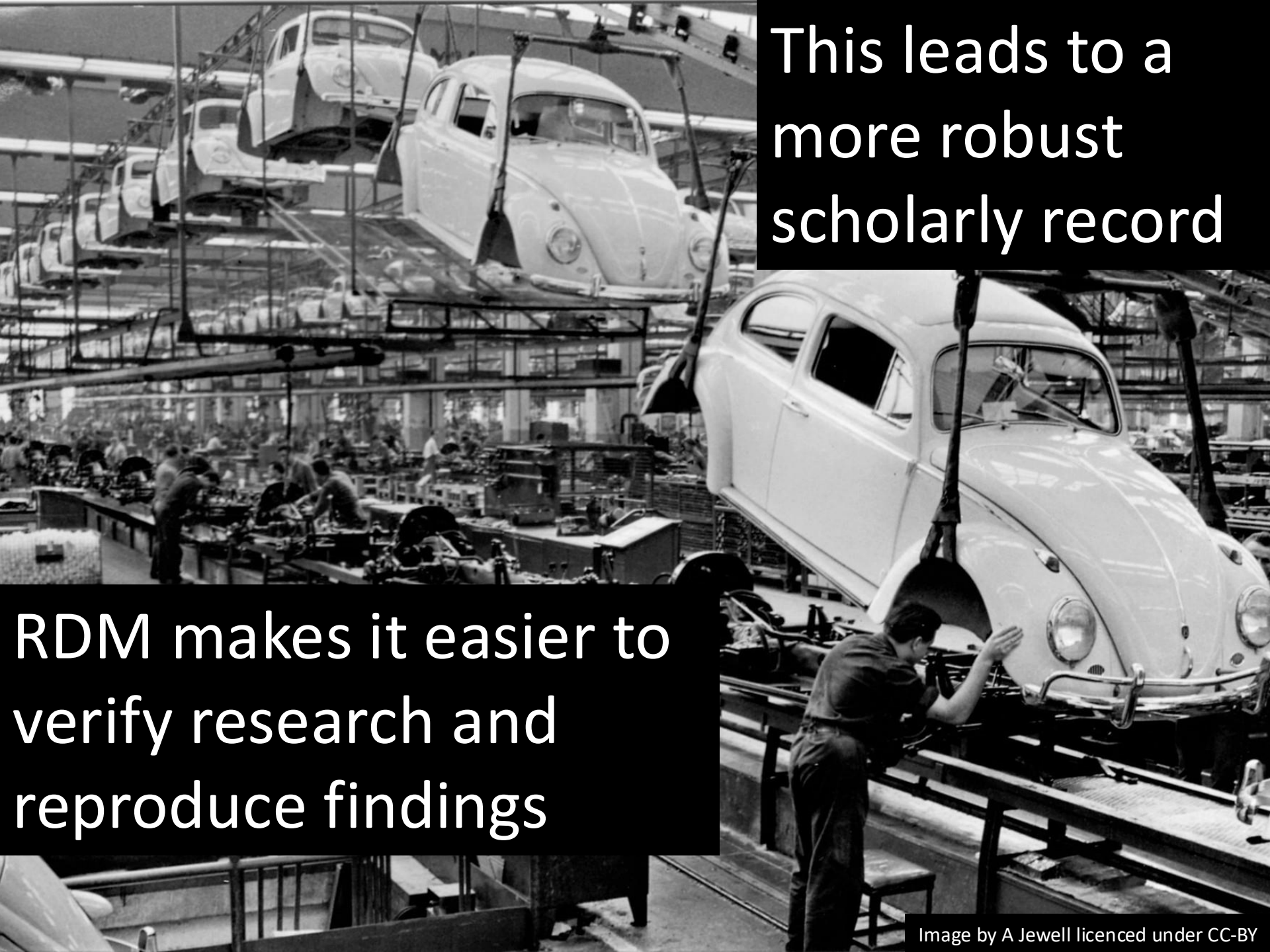
Part one:

RDM CONCEPTS AND TERMINOLOGY

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

RDM is a quality issue


RDM leads to increased transparency of the research process
Research data
Code
Methods
Documentation



This leads to a more robust scholarly record

RDM makes it easier to verify research and reproduce findings

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

Common 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


TERMINOLOGY

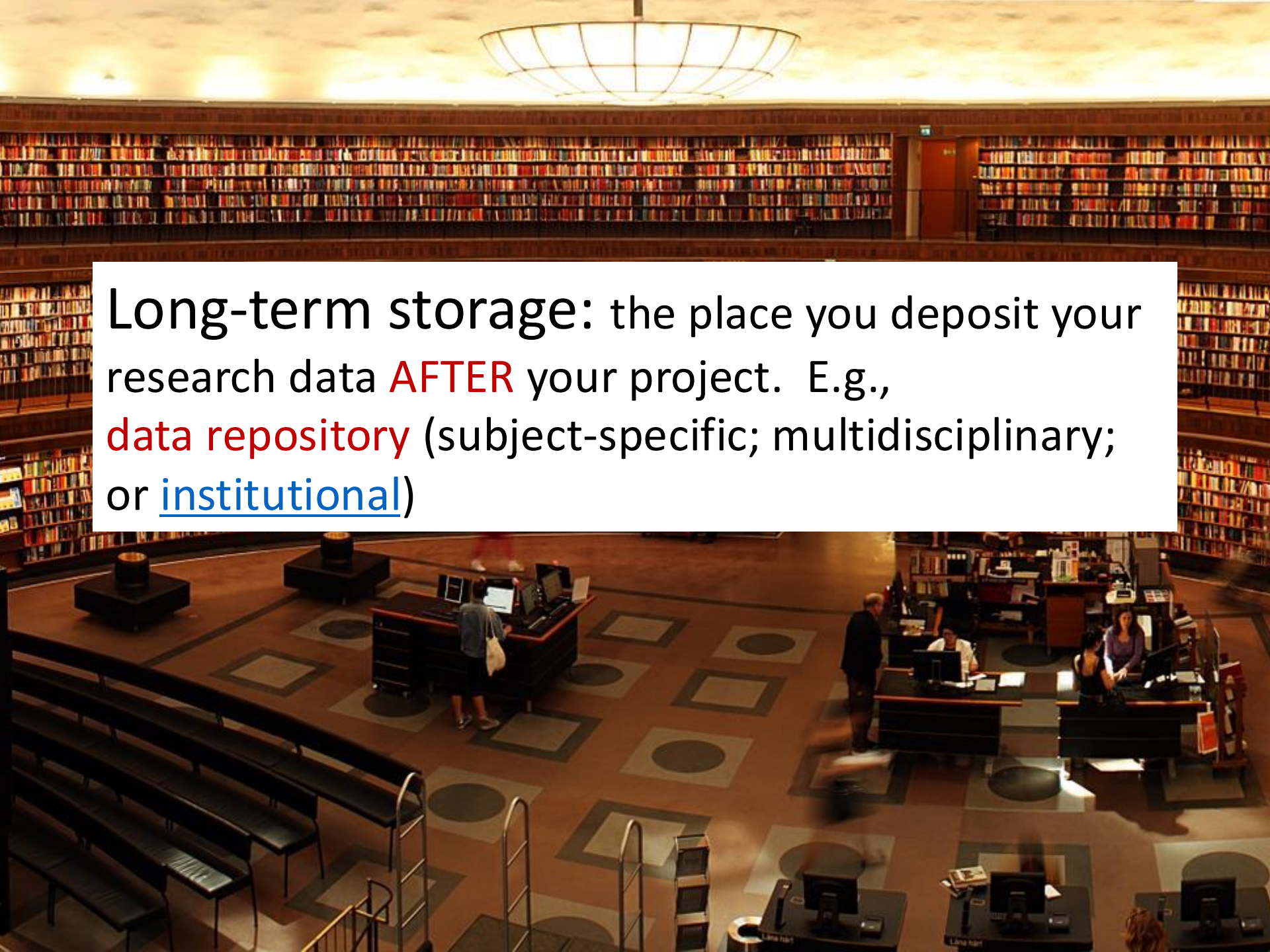
Anonymisation:

the art, or craft, of lowering or minimising the risk of individuals being identified from data

Source: https://ukanon.net/

Research on re-identification:
https://cpg.doc.ic.ac.uk/individual-risk


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)

Digital preservation:

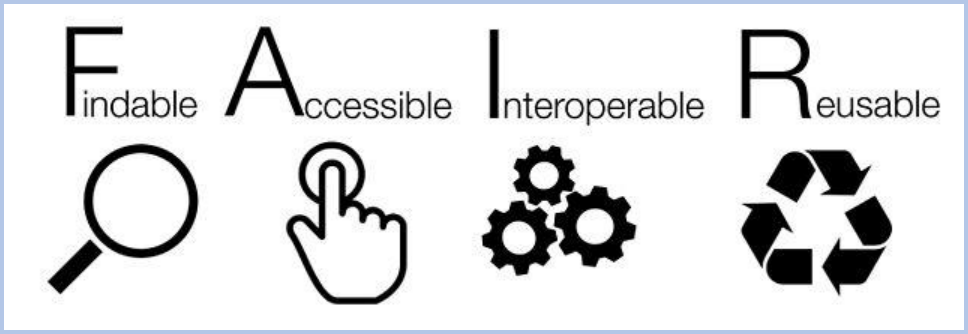
the series of managed activities necessary to ensure continued access to digital materials for as long as necessary

Data sharing has two meanings:

1. Post-research: sharing by publishing research data in a data repository
2. During research: sharing data with others

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**

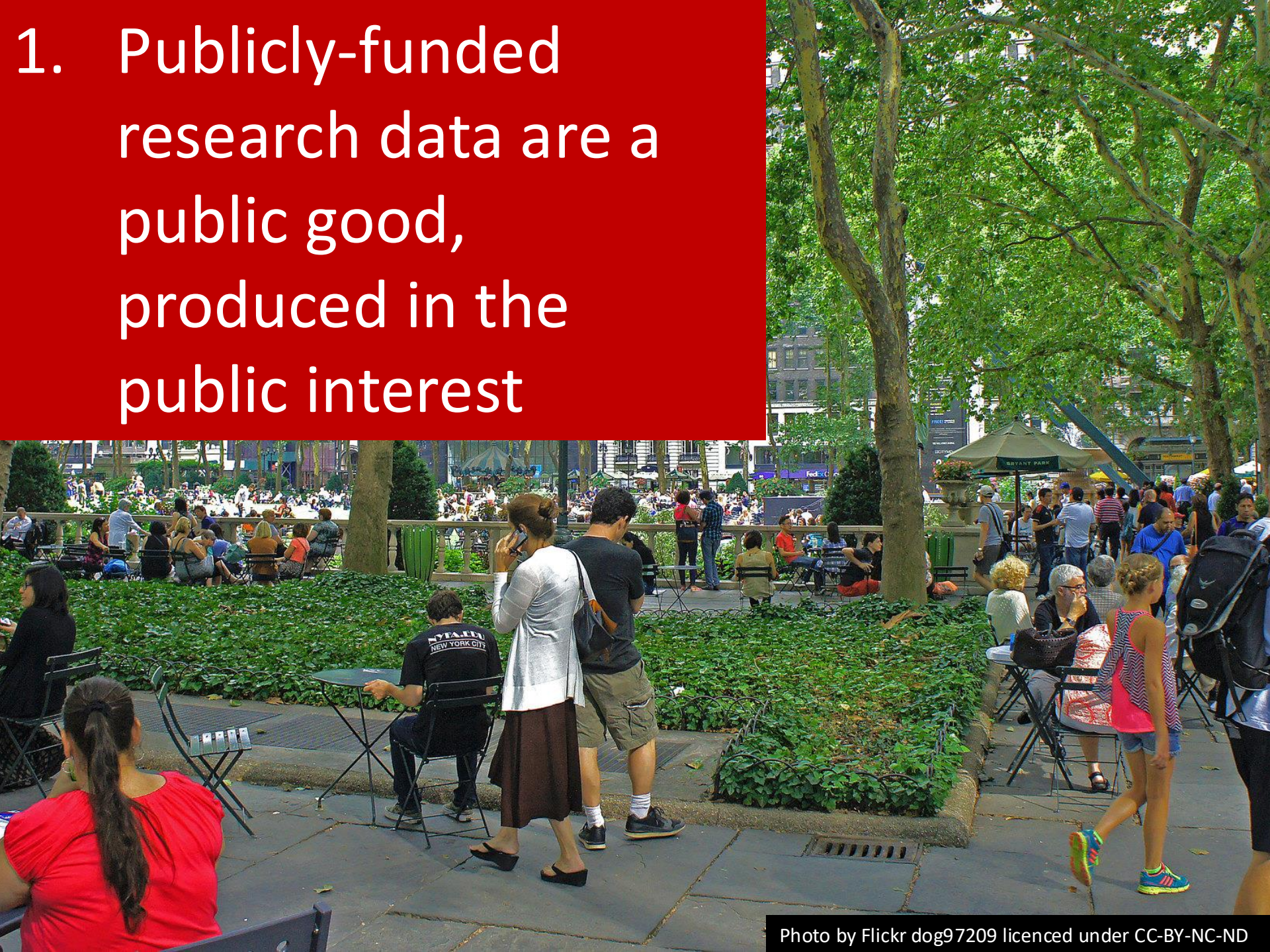
as opposed to unFAIR data

Part two:

DATA MANAGEMENT PLANNING


UKRI COMMON PRINCIPLES
ON DATA POLICY

1. Publicly-funded research data are a public good, produced in the public interest



2. Publicly-funded research data should be openly available to the maximum extent possible


Plan to keep your sensitive data safe
ABUS
65/50
ABUS


Plan to maximise the re-use potential of your data

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

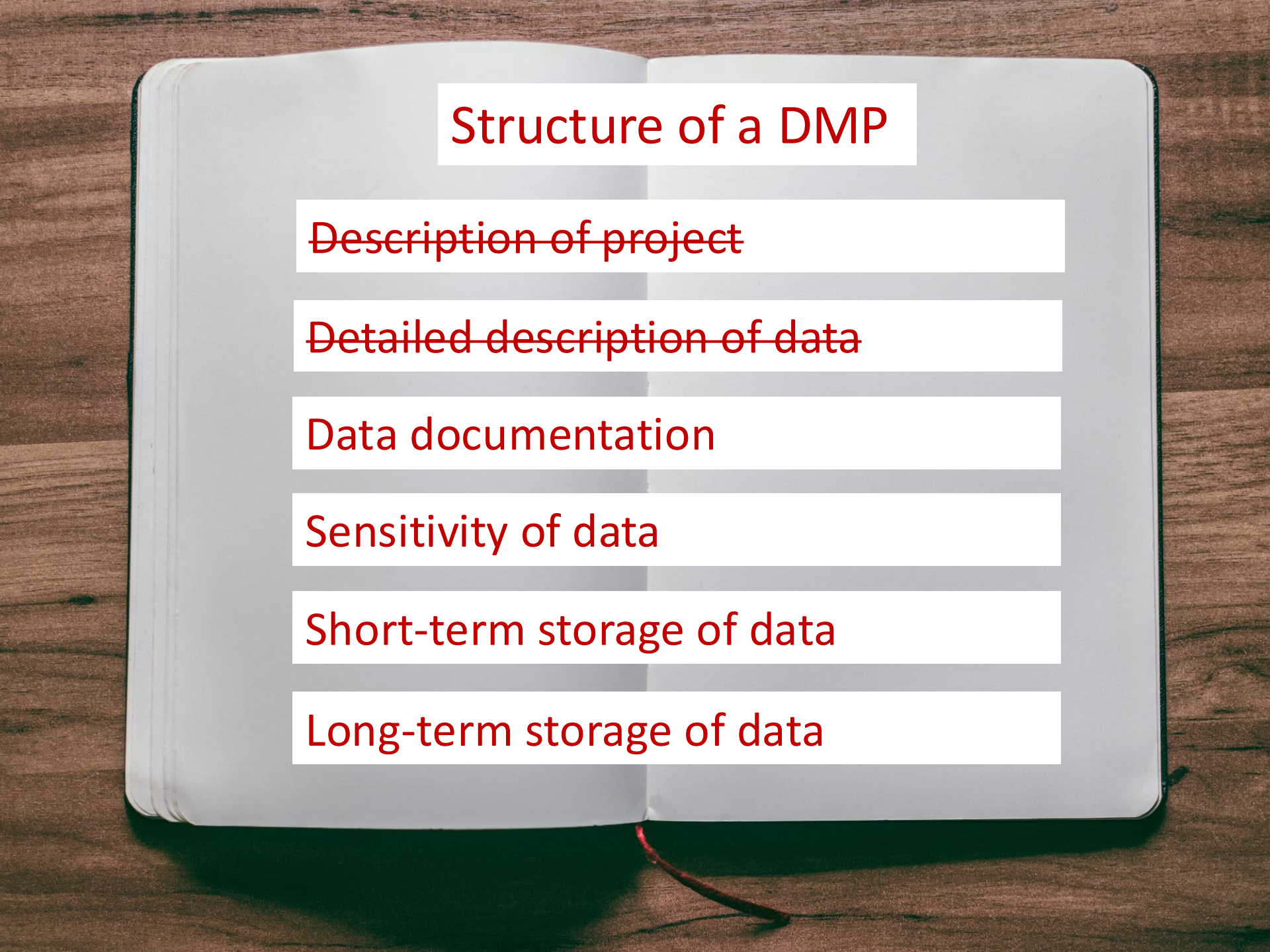
Data created or collected	Data type	Data format	Volume or duration	Planned storage or access
Raw ethnographic field notes	Notebooks	Paper	n/a	No shared access
Photographs	Digital images	JPEG	100 photographs * 4 Mb each = ca. 400 Mb	OneDrive for Business
Interviews	Sound recordings	MP3	20 interviews, 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 Durham research data repository
Magma flows	X-ray images	Tiff	50 Tb storage per year for five years = 250 Tb total storage	Shared Research Storage

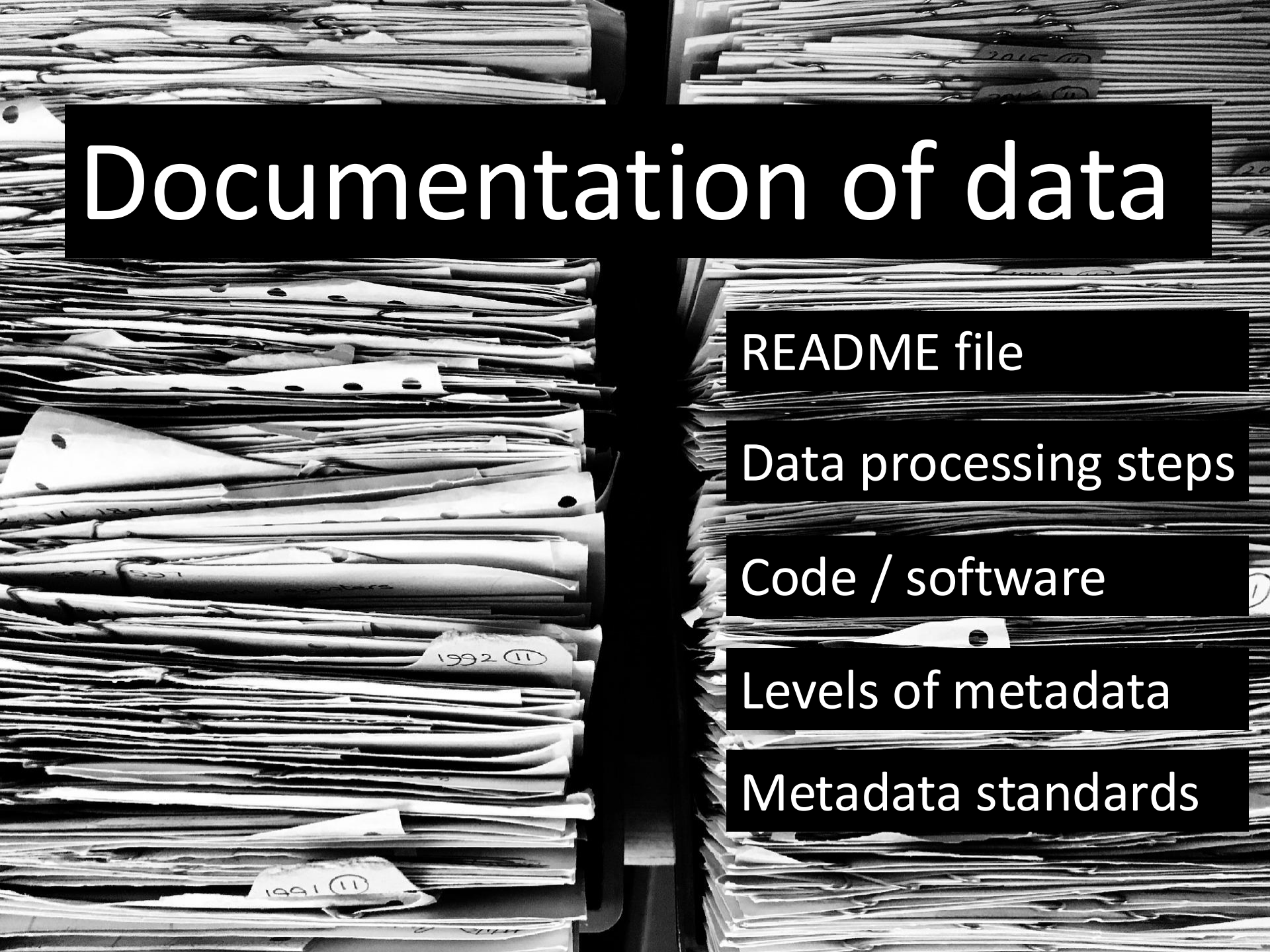
Example row in table

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	Shared Research Storage

Task one:

DESCRIBE YOUR RESEARCH DATA IN A TABLE


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



Documentation of data

- README file
- Data processing steps
- Code / software
- Levels of metadata
- Metadata standards

Sensitivity of data (from least to most sensitive)

Non-personal data (least sensitive)

Anonymised data

Pseudonymised data

Personal data

Special categories of personal data:

race	genetic data
ethnic origin	biometric data
political opinions	health data
religious or philosophical beliefs	sex life
trade union membership	sexual orientation

Secret data (most sensitive)

Short-term storage options

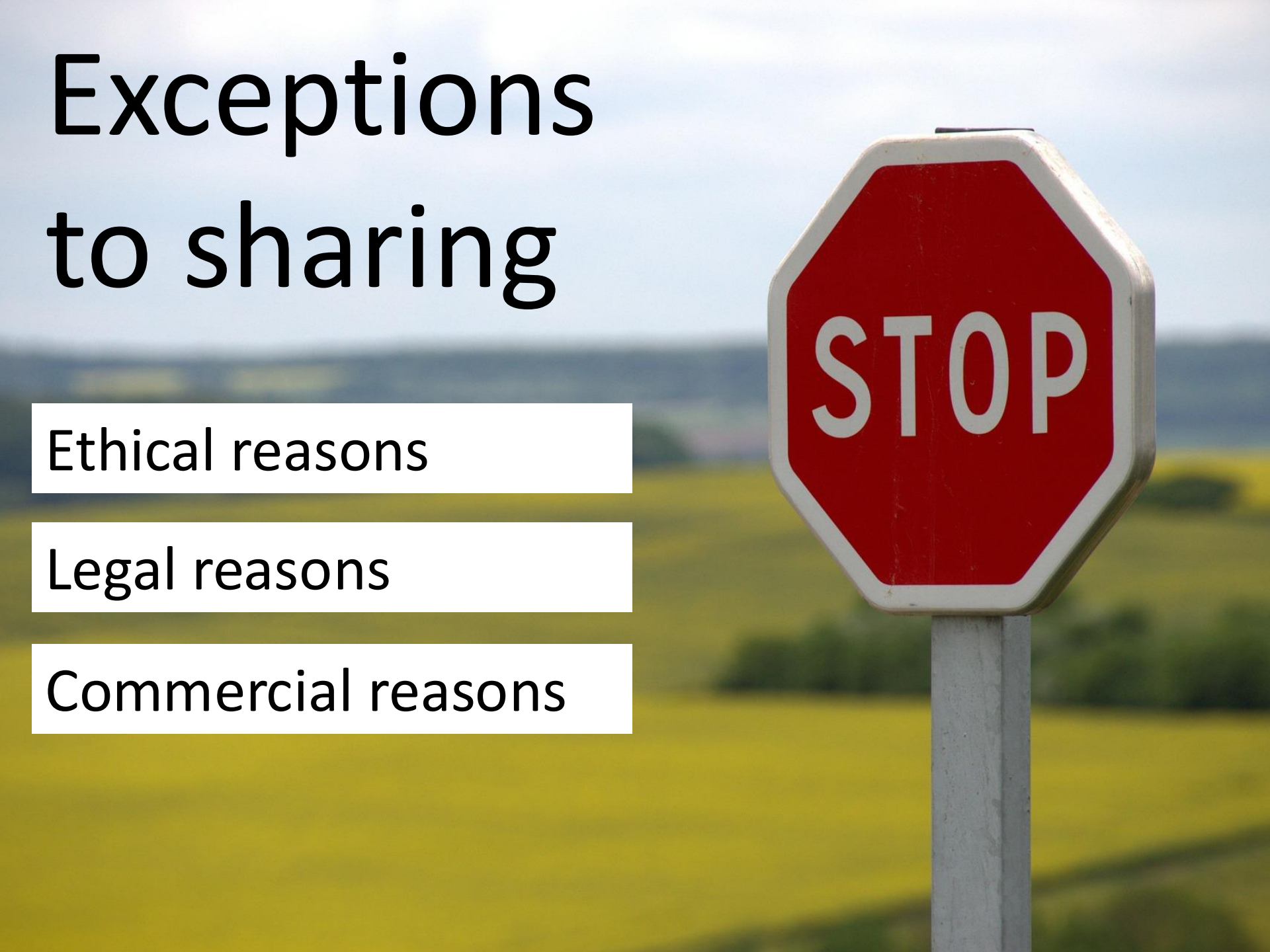
OneDrive for Business	Personal Research Storage	Shared Research Storage
1 TB encrypted cloud storage	1 TB storage on site (not encrypted)	Intended for funded research groups.
Easy to share with external collaborators	Complicated to share with external collaborators	Complicated to share with external collaborators
Automatic replication	Storage tier: Silver 3 Performance level: Silver Protection level: 3	Different storage tiers available
Managed with file explorer	Managed with CIS Storage Manager	Managed with CIS Storage Manager

Long-term storage options

Subject-specific repository	Multi-disciplinary repository	University data repository
The best option if you can find a suitable repository.	Zenodo Open Science Framework	collections.durham.ac.uk
Search for a repository using: FAIRsharing website	Dryad Mendeley	Personalised support for all research-active staff and students
e.g., ReShare for economic and social datasets	figshare	700 GB storage

Guidance on choosing a repository.

Exceptions to sharing



- Ethical reasons
- Legal reasons
- Commercial reasons

Do's and don'ts of writing a good DMP

DO	DON'T
Include as much detail as possible about your plans. Name specific tools / software you will use.	Assume your DMP reviewer knows what you mean. Explain what you mean in detail.
Use a DMP template.	Invent a new format for a DMP.
Describe all the research data you will create, preferably in a table.	Confuse research data with other research outputs. Focus on research data only.
Ensure all parts of your DMP work together coherently.	Contradict yourself. Skim your draft DMP in order to eliminate contradictions.
Try to answer all questions in your DMP template.	Be afraid to ask for help. This will exist both within institutions, and via national / European support organisations

Part three:

DEMONSTRATION OF DMPonline

DMPONLINE My Dashboard Create plans Reference Help Language nicholas syrotiuk

Durham University

1. Choose funder template

2. Write detailed plan

3. Share plan

Public DMPs

Public DMPs are plans created using funder guidelines.

Project Title				
Measuring understanding of biological data models.	University of Manchester Generic Template	University of Manchester	Yo Yehudi	
Charters for Better Work Better Lives: An Indian Partner Network	University of Manchester Generic Template	University of Manchester	Wendy Olsen	
Pentoxifylline for sepsis in preterm infants	Data management ZonMw-template 2019	Other	sinno simons	
Identifying Mechanisms of Failure of Cartilage Repair Surgery	University of Manchester Generic Template	University of Manchester	Gwenllian Tawy	
Simulating tourism water consumption with stakeholders (SIMTWIST)	Data Management Plan NWO (English)	Wageningen University and Research Centre (Netherlands)	Bas Amelung	

Business for Peace: Conflict Prevention and the Private Sector in South and Central Asia

Project Details | Plan overview | Write Plan | Share | Download

expand all | collapse all

16/16 answered

Data Summary (1 / 1)

Data Collection (1 / 1)

Short-term Data Storage (2 / 2)

Guidance

3. How will the data be stored in the short term?

Data collected by commissioned projects will be stored in secure locations - paper datasets will be stored in locked filing cabinets in secure offices, and electronic data on encrypted harddrives. Electronic data will also be backed up to Durham University's GDPR-compliant Box cloud.

Save

Answered 1 month ago by nicholas.syrotiuk@durham.ac.uk

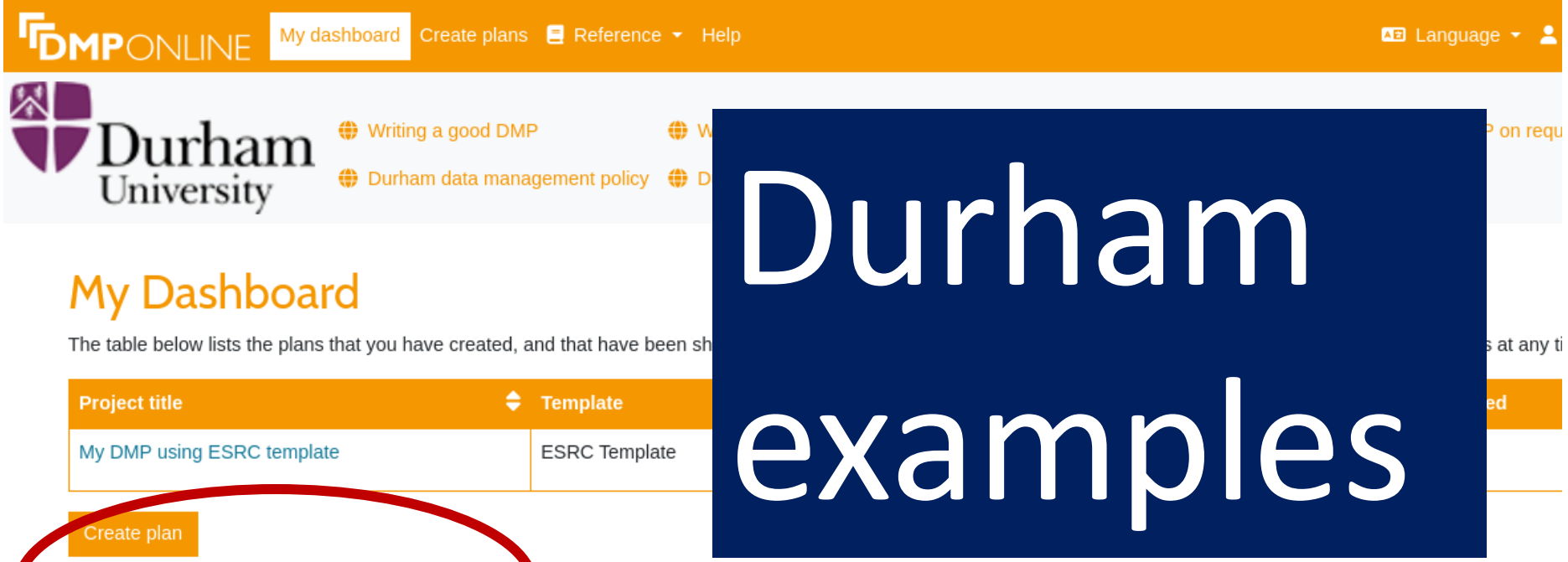
Guidance | Comments

AHRC | DCC | DU

Storage & security

Please familiarise yourself with the University's infomation security web pages. In particular you should become familiiar with Durham's Information Security Classsification and Handling Standard.

If you are planning to use **Shared Research Storage (SRS),** please include an estimate for research data storage costs in your grant


DMPONLINE
My dashboard
Create plans
Reference
Help
Language
Durham University
Writing a good DMP
Durham data management policy
Durham examples
My Dashboard
The table below lists the plans that you have created, and that have been sh
Project title
Template
My DMP using ESRC template
ESRC Template
Create plan
Durham University's Plans
The table below lists the plans that users at your organisation have created and shared within your organisation. This allows you to download a PDF and view their plans as samp
Search
Project title
Template
Owner
Updated
Writing a good data management plan (...
ESRC Template
nicholas.syrotiuk@durham.ac.uk
09-01-2023
'Wealthmaxxing your way out of loneli...
DU Standard RDM Template: Version one
fqmf28@durham.ac.uk
31-01-2022
Understanding Popular Xenophobia in C...
AHRC Data Management Plan
jonathan.saha@durham.ac.uk
18-08-2022
Treaty Scrutiny: The Role of Parliame...
ESRC Template
holger.hestermeyer@kcl.ac.uk
02-11-2022
Transparency and Judicial Review: An ...
DU Standard RDM Template: Version one
elizabeth.a.o'loughlin@durham.ac.uk
11-10-2022
Translation and (Counter)Espionage: A...
AHRC Data Management Plan
sergey.tyulenev@durham.ac.uk
07-03-2022

DMP REVIEW


Live
Demonstration

Part four: **Try writing a**

DATA MANAGEMENT PLAN (Task 2)

DMPONLINE | Home | Public DMPs | Funder requirements | Help | Language

Plan to make data work for you

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


Durham University
Library and Collections
Library Research Support
University Library and Collections / Guides / Library Research Support: Ope
Sear
Library Research Support: Open Research: Research Data Management
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
Research Publications Policy & Rights Retention
Paying for Open Access
Publisher Deals
Durham Research Online (DRO)
REF OA
PGRs & eTheses
Research Data Management
Open Access Week 2023
Monthly Activity & News
SafePod: Secure Data Access
What do we mean by Research Data Management?
Our researchers create, collect and measure large amounts of research data every day, and some of this data is used to support our
https://libguides.durham.ac.uk/open_research/rdm
spirit of open scholarship, researchers are expected to demonstrate they are working transparently and reproducibly. Researche

Durham Research Methods Café

Every other Wednesday of term
All PGR/PGT students and DU staff Welcome

Conversation topic: When things go wrong!

28th February 2024, 1100-1230

Discussing what to do when research doesn't go as planned or appears to be failing.

Free Coffee/Tea and pastries!

Check out our Research Methods Café online for methods-related chat/announcements, and online access to the Café Conversations.

Location:
First Floor, Arthur Holmes Building: left of Calman Learning Centre, upstairs, turn left.
Door signed 'DRMC'.

Co-hosts:

- Advanced Research Computing (ARC),
- Durham Research Methods Centre (DRMC)
- Research Support, Durham Library


Thank you
Nicholas Syrotiuk
@DurhamRdm

Inkpath

Writing a good Data Management Plan (DMP)

From 07/11/2024 09:30 to 07/11/2024 11:00

--- PLEASE NOW MARK YOUR ATTENDANCE AT THIS ACTIVITY ---

Mobile App Users

1. Open the Inkpath App
2. Press this button
3. Scan this QR Code

4. Confirm attendance Complete

Web Users

1. Open **webapp.inkpath.co.uk**
2. Press this button Record Attendance
3. Enter this unique code

5863303521

4. Press SUBMIT
5. Confirm attendance Complete

References:

CESSDA Training Team (2017 - 2022). CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC. DOI: http://doi.org/10.5281/zenodo.3820472

Information Commissioner's Office. https://ico.org.uk/

Rocher, L., Hendrickx, J.M. & de Montjoye, YA (2019): "Estimating the success of re-identifications in incomplete datasets using generative models." Nat Commun 10, 3069. DOI: https://doi.org/10.1038/s41467-019-10933-3

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

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

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