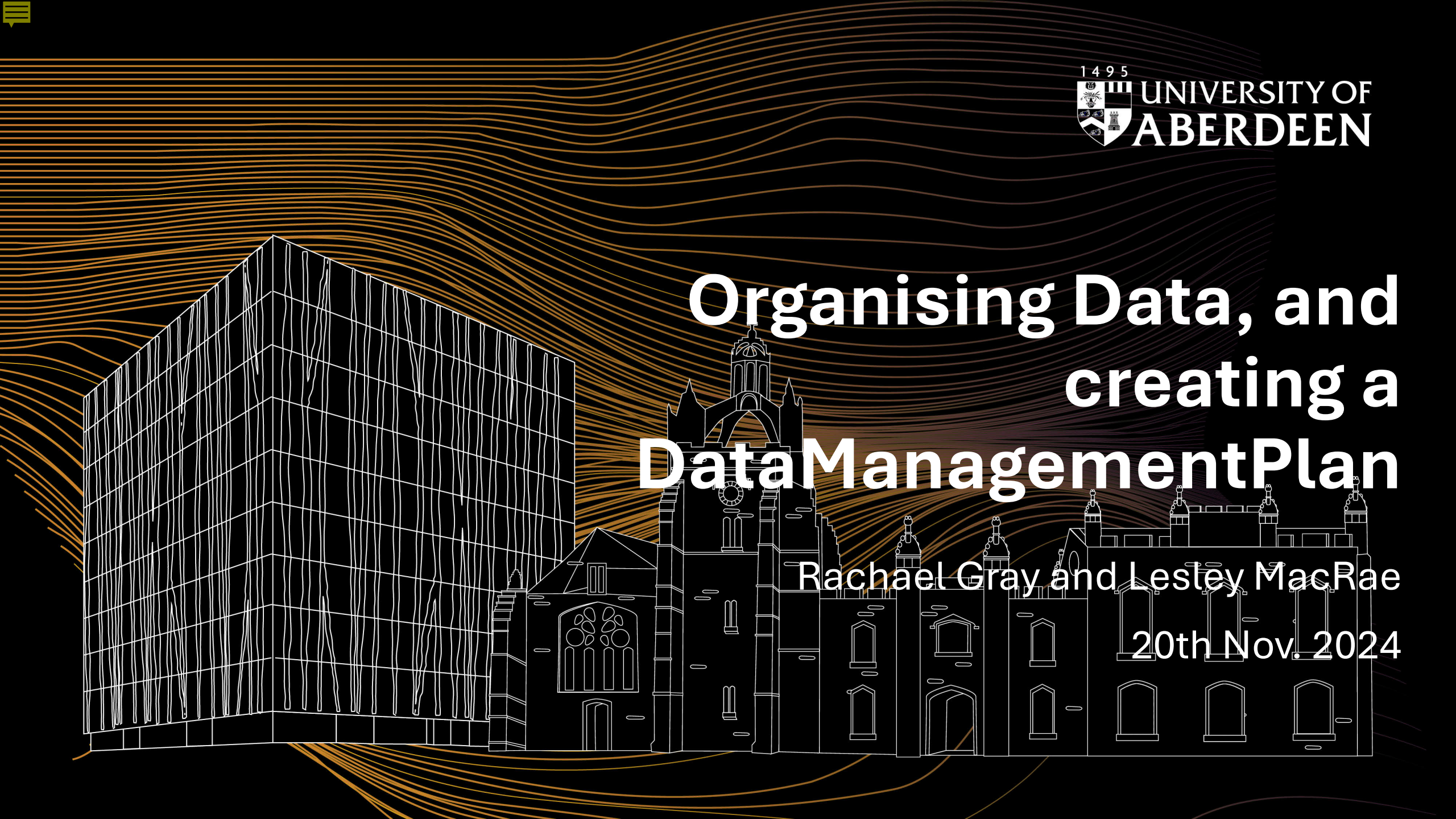


Organising Data, and creating a DataManagementPlan

Rachael Gray and Lesley MacRae

20th Nov. 2024

Overview

- Who are Digital Research Services?
- Before you start your DMP
- Describing your Data
- Data Collection/Generation
- Documentation and Metadata
- Ethical and Legal Compliance
- Storage and Backup
- Selection and preservation
- Data Sharing
- Data Sharing
- Responsibilities
- Policies
- Common Pitfalls
- Support and Resources
- Organising your data
- Repositories

Who are the Digital Research Team?

We are a team that sits in the Directorate of Digital and Information Services, with the know-how to provide bespoke technological solutions to your research needs.

We combine our knowledge of the research community with our technical expertise to provide the best solution for you.

We can be reached at digitalresearch@abdn.ac.uk

Digital Research | Research | The University of Aberdeen (abdn.ac.uk)

Who are the Digital Research Team?

Our services cover:

- Access, use, and support with the High-Performance Computer (HPC)
- Access and support for our survey and data entry tools
- 1-1 consultancy for research solutions
- DMP support
- Data management support
- Access to the Data Safe Haven and SafePod

Digital Research | Research | The University of Aberdeen (abdn.ac.uk)

What is a DMP? A quick definition

A DMP, or Data Management Plan, is a formal ‘living’ document that outlines how your data will be handled, including creation, maintenance and sharing, from the start of your project to after your project is complete.

Most funding bodies are now expecting to see a DMP and for it to be maintained throughout the project’s lifecycle.

Data Management Plans | DCC

Data Management Plan | University Systems and Software | Toolkit | The University of Aberdeen (abdn.ac.uk)

A brief outline of data protection laws

Data Protection Act (2018)

Any use of Personal Data should be

- used fairly, lawfully and transparently
- used for specified, explicit purposes
- used in a way that is adequate, relevant and limited to only what is necessary
- accurate and, where necessary, kept up to date
- kept for no longer than is necessary
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage

Data Protection Act 2018 (legislation.gov.uk)

Data Protection | StaffNet | The University of Aberdeen (abdn.ac.uk)

A brief outline of data protection laws

General Data Protection Regulations (GDPR)

GDPR aims to protect individuals' privacy rights by:

- Granting individuals rights over their data, such as the right to deletion, correction, and access (where applicable)
- Requiring consent for the use of personal data
- Ensuring the fair, lawful and transparent use of personal data
- Mandating security measures to protect data, mandating the reporting of data breaches, and appointing data officers
- However, there are exemptions to these rights for researchers, but these need to be explicit to individuals

General Data Protection Regulation (GDPR) – Official Legal Text (gdpr-info.eu)

UK GDPR guidance and resources | ICO

Data Protection | StaffNet | The University of Aberdeen (abdn.ac.uk)

You may need to fill out a Data Protection Impact Assessment

If you plan on doing any of the following, you'll need a DPIA:

- use systematic and extensive profiling with significant effects on individuals
- process special category or criminal offence data on a large scale
- systematically monitoring publicly accessible places on a large scale
- use profiling or special category data to decide on access to services, opportunities or benefits
- profile individuals on a large scale
- match data or combine datasets from a different source
- profile children or target marketing or online services at them
- process data that might endanger the individual's physical health or safety in the event of a security breach

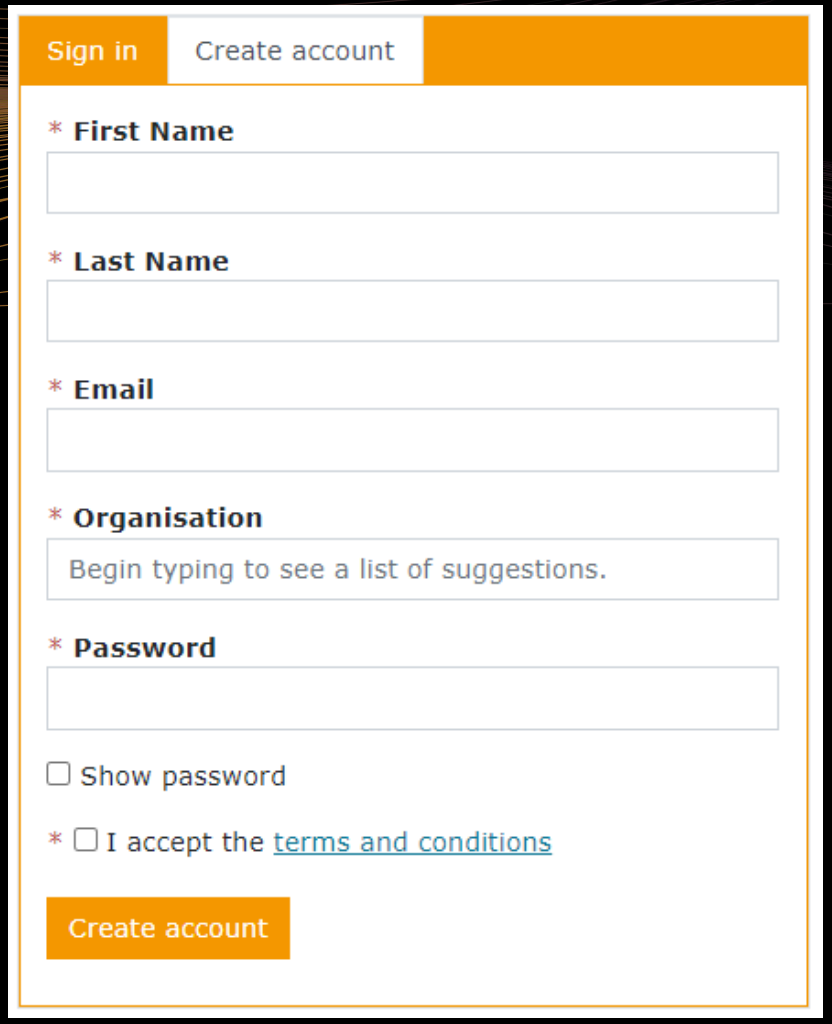
More information can be provided by the Information Governance team at dpa@abdn.ac.uk

Data Protection | StaffNet | The University of Aberdeen (abdn.ac.uk)

Where can a DMP be created?

The university has access to a DMP tool called DMPonline - DMPonline (dcc.ac.uk)

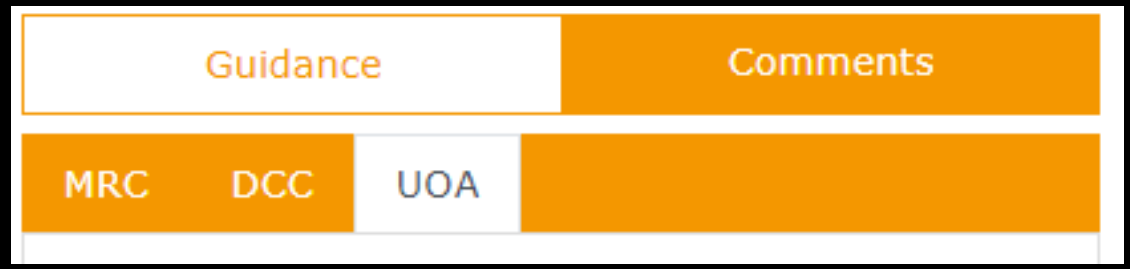
It’s free to use and you can start creating DMPs after you’ve created an account.



Different funders have slightly different requirements…

However, each DMP will contain the same key areas that will be covered in this presentation.

Each area will also contain descriptions of what info should be contained – these descriptions are also geared to which funder you have submitted.



Funders' data plan requirements | DCC

Describing your Data

What type of study are you partaking in?

Prospective

Retrospective

Case-Control

Please note, that if you're using or re-using any third-party data, make sure you have the appropriate permissions to allow for data sharing and/or preservation plans

[DMP 1.1]

What data will you collect or create?

Qualitative	Computational Models	Statistics	Measurements
Text	Images	Audio Visual Data	Samples

Please note, that if you're using or re-using any third-party data, make sure you have the appropriate permissions to allow for data sharing and/or preservation plans

[DMP 1.2]

How much data and what format?

Describe what format the data will take, such as .csv, .txt etc and why you chose this format.

How much will cover how much data storage you will need for your project; Will you only require a couple of gigabytes or multiple terabytes?

[DMP 1.3]

Recommended formats — UK Data Service


Data Collection/Generation

How will you collect/generate your data?

Here, you will describe how your data will be collected, generated and/or processed, with what software and why

Excel/Other spreadsheet software

SPSS/Other statistical analysis programs

Surveys/Survey tools

Audio/transcriptions/video

[DMP 2.1]

Please be aware if you're using third-party apps or software

You may need to fill out a Supplier Cyber and Data Assessment

IS18: Supplier Cyber and Data Assessment Process Overview (abdn.ac.uk)

How will you ensure data quality?

Data Coding

Data Cleaning

Repeat Samples

Validation

Standardisation

Data Management, Documentation and Curation

How will you manage, store, and curate the data?

C DRIVE

H DRIVE

SHARED DRIVE

USB/ EXTERNAL HARD DRIVE

CLOUD

[DMP 3.1]

Research data management — UK Data Service

Data Storage for Researchers | University Systems and Software | Toolkit | The University of Aberdeen (abdn.ac.uk)

But please note this…

External storage can be fragile in many ways

Aberdeen university students evacuated following bin fire in classroom

Evening Express

Fire sparks evacuation at Aberdee

Fire sparks evacuation at Aberdeen university

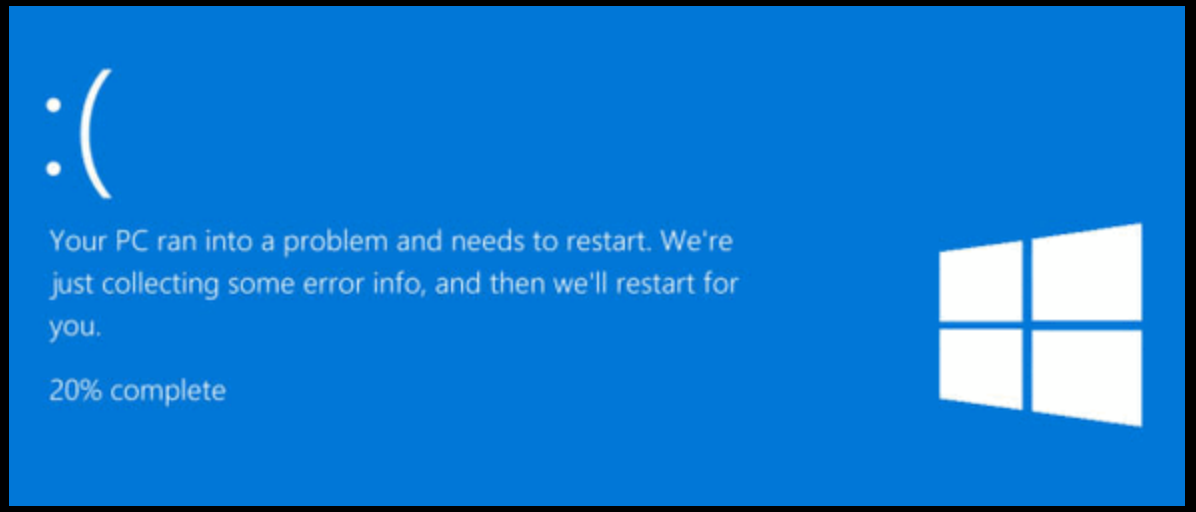
by Anthony Joseph 20/11/2014, 9:57 am

Aberdeen university students evacuated following bin fire in classroom (pressandjournal.co.uk)

Aberdeen students evacuated from fire (pressandjournal.co.uk)

But please note this…

- Your H drive should only be for personal work and not for your research data
- Your C drive should only be for transient storage, and anything saved here needs to be backed up on a shared drive or your OneDrive



Data Storage & Archiving | StaffNet | The University of Aberdeen (abdn.ac.uk)

What are the main risks to your data and how to safeguard?

- You will also need to state the risks to your data and why the risks are there
- What policies are in place to mitigate this risk?
- What formal standards will you implement to mitigate risk?

[DMP 3.1]

Data Protection | StaffNet | The University of Aberdeen (abdn.ac.uk)

What metadata and documentation will you also produce?

- Metadata describes your data, how it was produced, with what procedures and why
- It describes and contextualises your data to make it accessible and understandable to others
- You will also need to think about metadata standards:


Descriptive


Technical


Administrative


User


Preservation

[DMP 3.2]

What are Metadata Standards | DCC

What is your data preservation strategy?

Here will be a brief description of what data you will keep, for how long, where, and what policies or standards you will follow.

Time

Software upgrades/decommission

Data you'll destroy

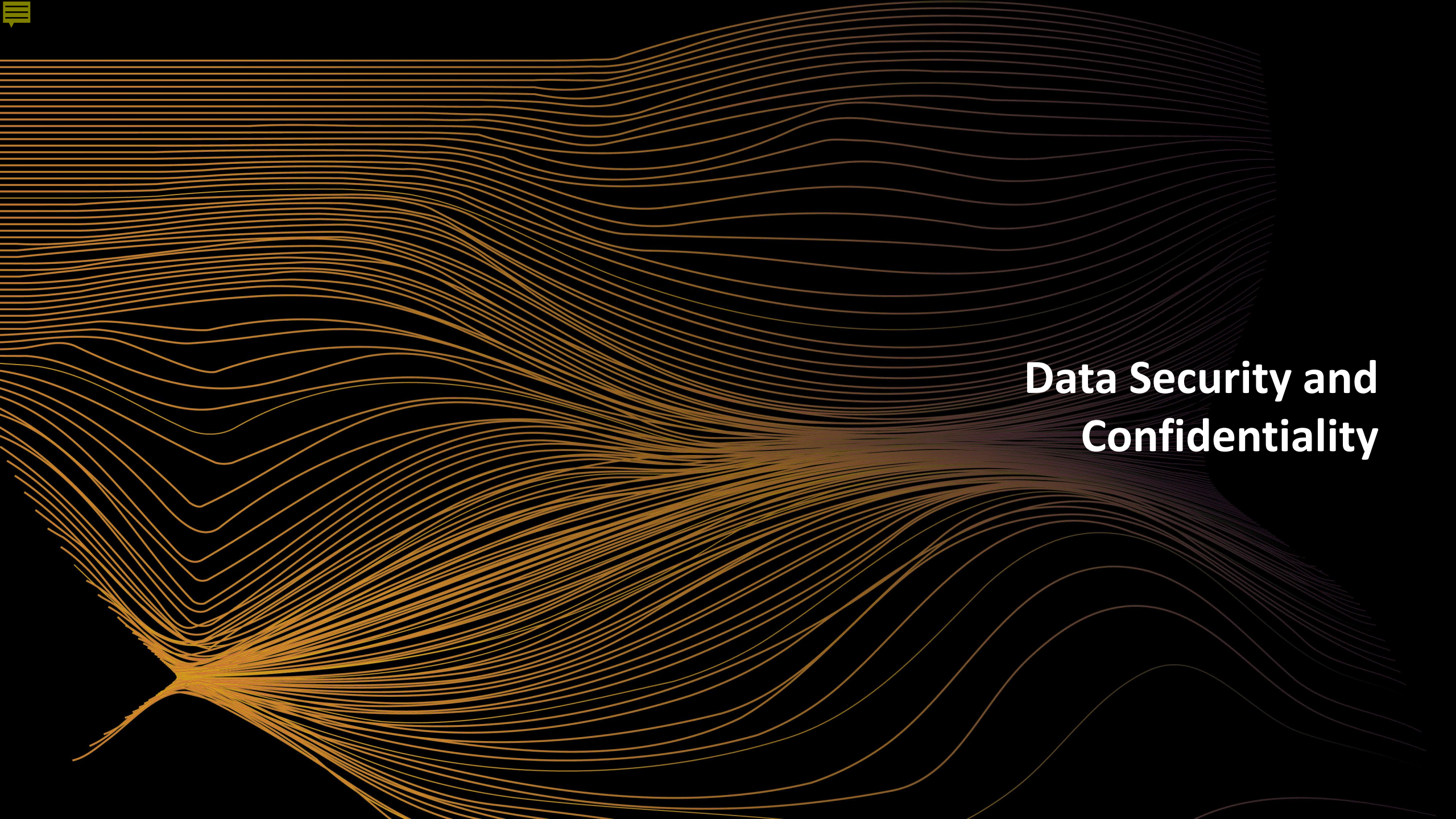
Archiving

Policies/Standards

[DMP 3.3]

Research Data Management policy.pdf (abdn.ac.uk)

research_data_management guidance.pdf (abdn.ac.uk)



Data Security and Confidentiality

How will you protect your data?

You will need to consider the security policies and standards you will implement, the data storage method you'll use, who will have access, and how you will mitigate risk.

You will also need to outline the main risks to your data, such as unauthorised access, corruption, deletion, etc

[DMP 4.1/4.2]

Data Storage & Archiving | StaffNet | The University of Aberdeen (abdn.ac.uk)

Information Security Policy (abdn.ac.uk)

Information Security Policy (abdn.ac.uk)

Data Protection | StaffNet | The University of Aberdeen (abdn.ac.uk)


Data Sharing and Access

Suitability for sharing?

You will need to think about how and when you can share your data, whether you can share your data, and if you have the right permissions to share.

Most likely no, not yet

Think about consent

Exclusivity period

[DMP 5.1/5.4/5.5]

Research Data Management policy.pdf (abdn.ac.uk)

research_data_management guidance.pdf (abdn.ac.uk)

Discoverability of your data

Here, you'll describe where you will deposit your data, the cost of preservation of your data, and what data will be preserved.

WEB

PRESENTATIONS

PURE

WORD OF MOUTH

[DMP 5.2]

DataCite Commons

Who will govern access to your data?

PI

Steering Committee

Application form

Monitoring Publications

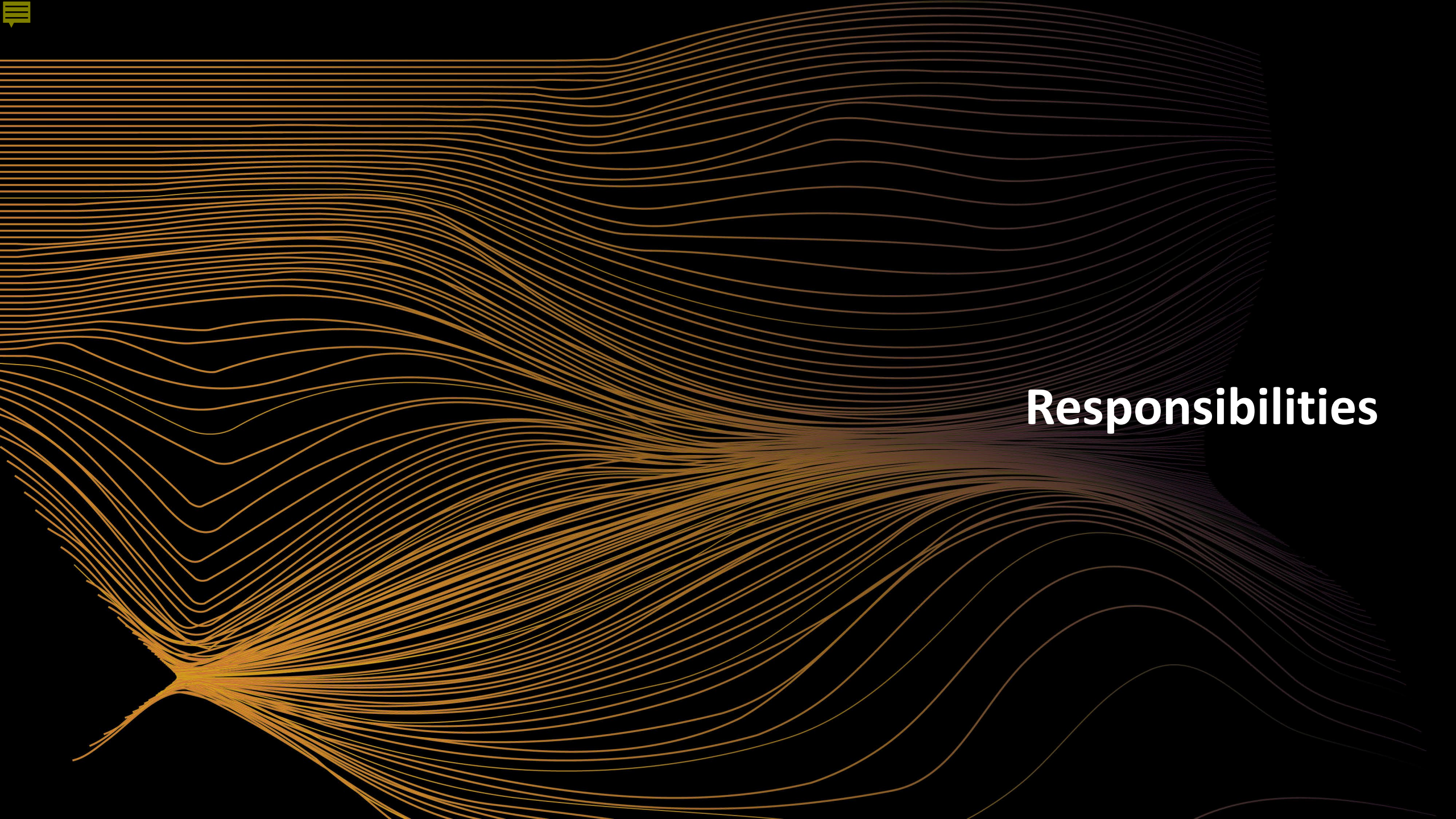
[DMP 5.3]

Regulation of responsibilities

You will need to think about whether you will need a Data Sharing Agreement for those wanting to access your data.

The short answer will most likely be yes.

[DMP 5.6]



Responsibilities

Who will be responsible for what?

You will need to outline who will be responsible for what role or activity and how you will coordinate collaborative projects.

Ultimately, as the PI, you are responsible for what happens in your project.

Examples of roles include:

- Data capture
- Metadata creation
- Quality assurance
- Storage and backup
- Data archiving
- Data sharing
- Plus, more….

[DMP 6]

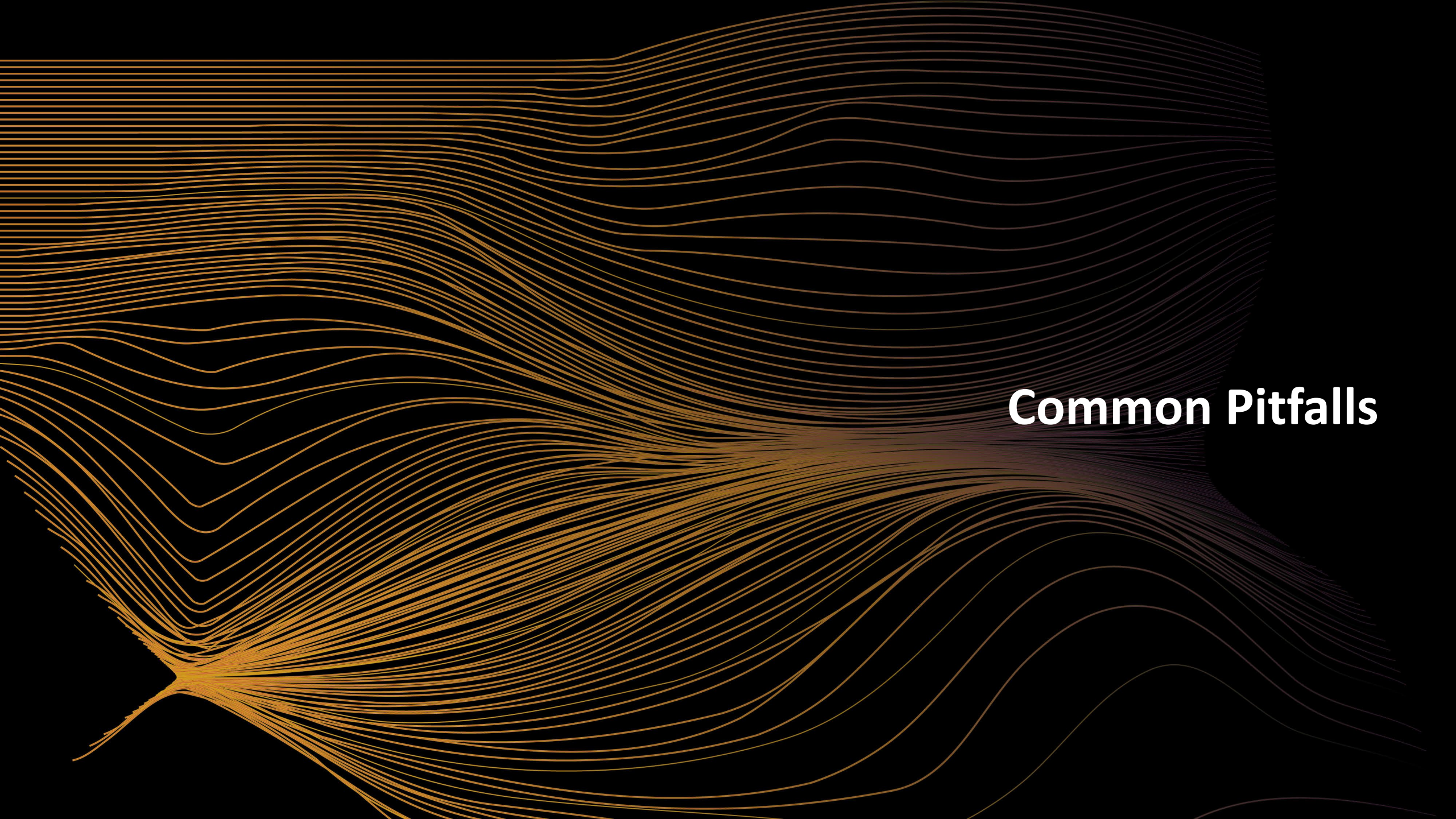
Relevant Policies

Include the relevant institutional, departmental, and study policies you will follow

All our institutional policies can be found in the Policy Zone

[DMP 7]

Search Policy Zone | StaffNet | The University of Aberdeen (abdn.ac.uk)



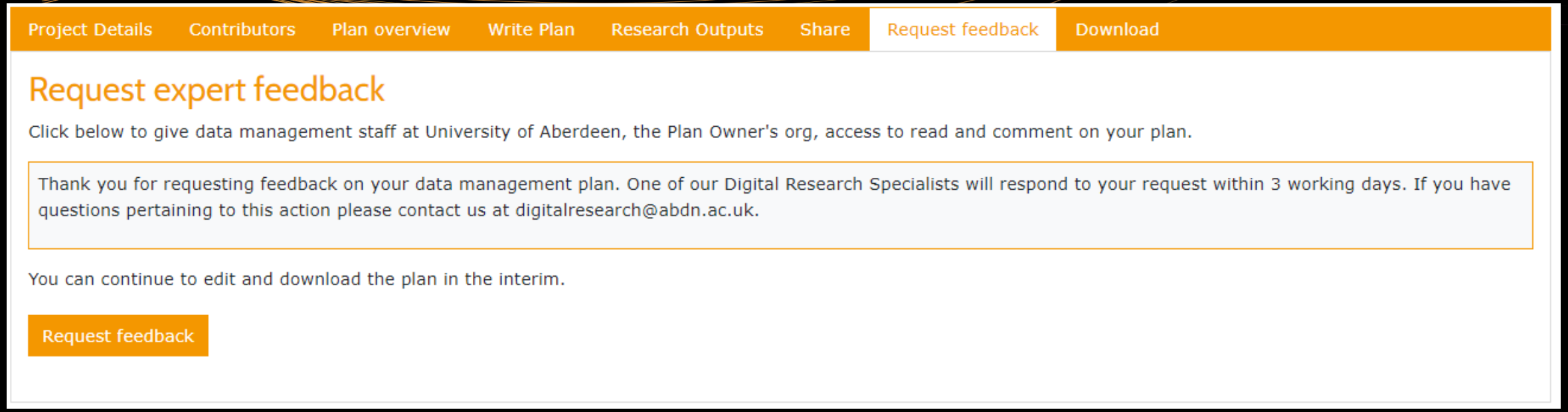
Common Pitfalls

Just to outline the common pitfalls we see

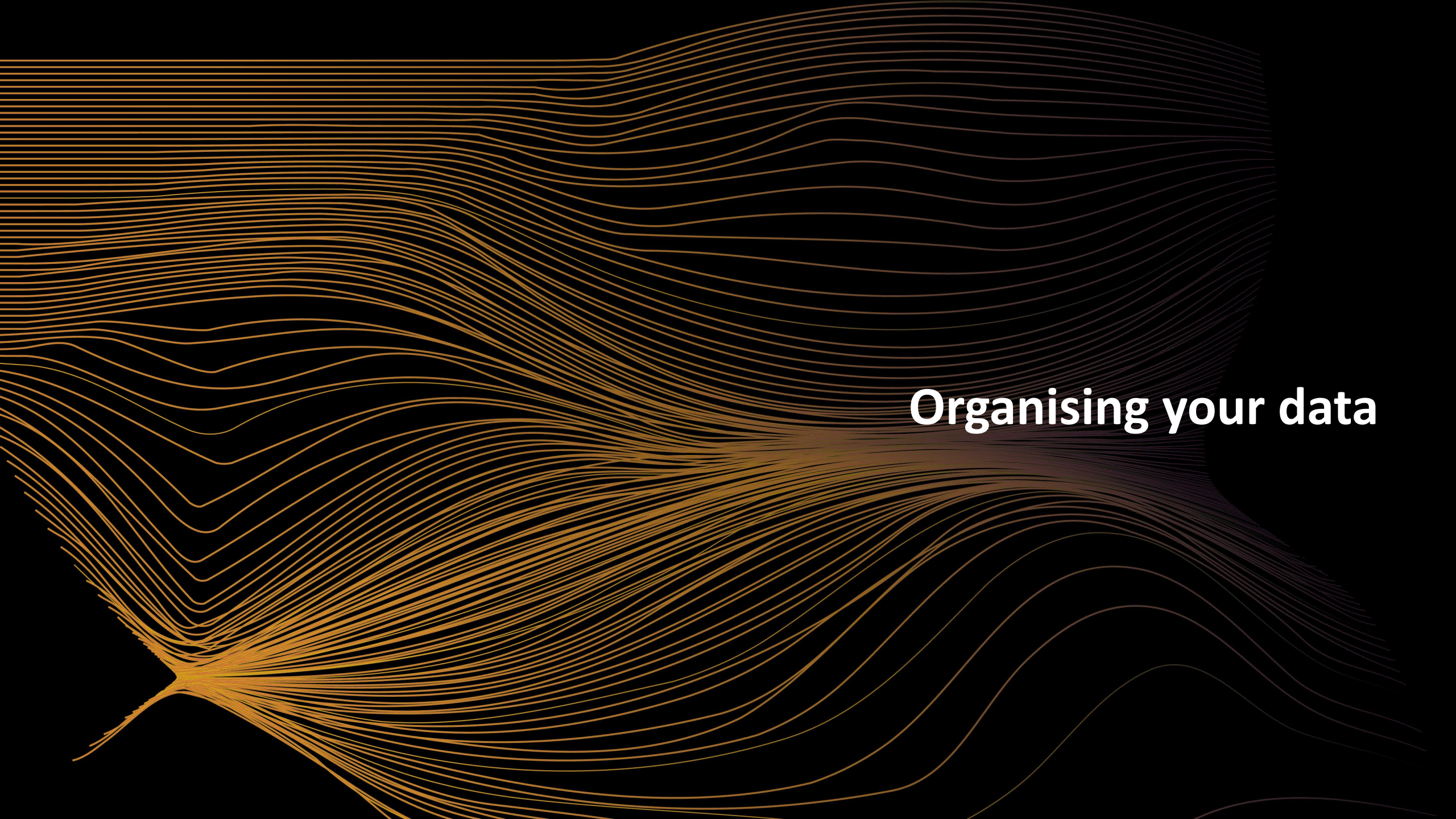
- Type, format and volume of data
- Who owns the data?
- How the data will be backed up, secured and stored?
- Who will you share the data with and under what conditions?
- Will you share the data in a repository, how you will handle access requests, or if you will use another mechanism?

Support and Resources

When you've started to create your DMP…



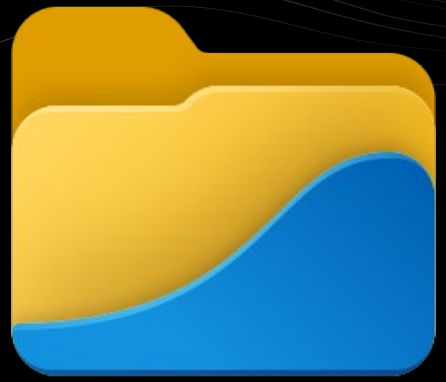
Data Management Plans | DCC



Organising your data

Organise your data

- Make folder names meaningful
- Hierarchical structure
- Store sensitive data in separate folders
- Check with colleagues – is there a scheme to follow?



Readme file : a how-to navigate the file structure (include any abbreviations)

File naming

- Develop a naming convention
- Be consistent!
- Avoid spaces, use_underscores, avoid special characters or symbols
- YYYYMMDD good format for dates

20170205_ProjectA_Ex1Test1_SmithE_v1.xlsx

Version control

- Avoids working on out of date versions
- Version naming : v.1,v.2,v1-0, v1-1
- First number for major changes, second for minor

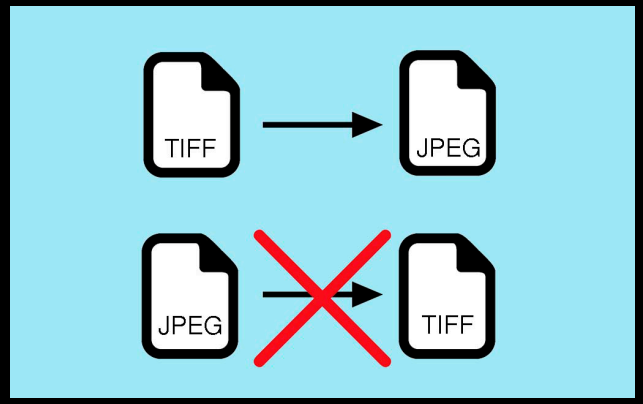
Tables : Use a control table, record the version no., date of change, person making it, nature of change

Git/Subversion : Useful for tracking changes to source code

File formats

- Formats at least risk of obsolescence
- Open, non-proprietary formats
- More widely used the better
- Keep original file formats (Word easier to preserve than PDF)

Webpage guidance.docx (sharepoint.com)



How will you manage, store, and curate the data?

- Writing documentation for your data will help you understand it – and others interpret/reuse it

- Disciplinary metadata : https://rdamsc.bath.ac.uk/

Examples of data documentation :

dataset structure
lab notebooks
protocols
codebook
methodology
interview schedules
software/computer code

PROJECT LEVEL

FILE LEVEL

VARIABLE LEVEL

[DMP 3.1]

Research data management — UK Data Service

When :

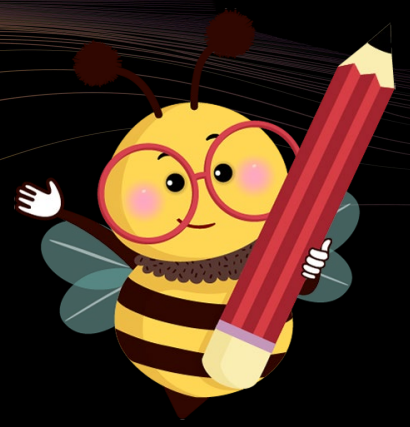
- Easier to do when you’re creating and it’s fresh in your mind

What :

- Who created the data, how it was gathered and for what purpose

Where :

- Readme, plain text file with readme ‘filename’ https://data.research.cornell.edu/data-management/sharing/readme/
- Information in the file itself (extra Excel sheet)
- Store in the folder with the dataset



Make a README

Because no one can read your mind (yet)

Data sharing in Repositories

Repositories

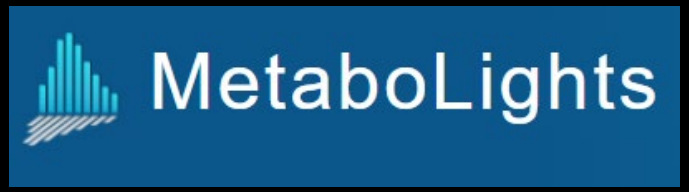
Generic Repositories :

- Accessible to a wide audience
- Interdisciplinary
- Simple metadata

Generalist-Data-Repository-Grid.pdf (agu.org)

Disciplinary Repositories

- Offer expertise
- Likely to accept complete datasets
- May be more selective in the data they accept



Resources

Digital Research | Research | The University of Aberdeen (abdn.ac.uk)

Data Management Plans | DCC

Data Management Plan | University Systems and Software | Toolkit | The University of Aberdeen (abdn.ac.uk)

Data Protection Act 2018 (legislation.gov.uk)

Data Protection | StaffNet | The University of Aberdeen (abdn.ac.uk)

General Data Protection Regulation (GDPR) – Official Legal Text (gdpr-info.eu)

UK GDPR guidance and resources | ICO

DMPonline (dcc.ac.uk)

Funders' data plan requirements | DCC

Recommended formats — UK Data Service

IS18: Supplier Cyber and Data Assessment Process Overview (abdn.ac.uk)

Data Storage for Researchers | University Systems and Software | Toolkit | The University of Aberdeen (abdn.ac.uk)

Research data management — UK Data Service

Data Storage & Archiving | StaffNet | The University of Aberdeen (abdn.ac.uk)

What are Metadata Standards | DCC

Research Data Management policy.pdf (abdn.ac.uk)

research_data_management guidance.pdf (abdn.ac.uk)

DataCite Commons

Policy Zone | StaffNet | The University of Aberdeen (abdn.ac.uk)