

# PATTERN.

Empowering Open and Responsible  
Research and Innovation

## Session 2: Planning for FAIR: Introduction to RDM and DMPs

(Beginner Level)

Course: FAIR Research Data Management

PATTERN



Funded by  
the European Union

# About these sessions



## Each Session:

- Is part of a **5-part series**
- Includes around **2-3 hours of teaching and learning content**
- Is being **recorded**





# About these sessions



**Variety** of learning activities

**General and discipline-specific** learning

**Designed for researchers**



Photo by [Jon Tyson](#) on [Unsplash](#)

# About these sessions



Project work

The sessions  
are **flexible**



Photo by [John Schnobrich](#) on [Unsplash](#)



# About these sessions



**Choose in person or  
online learning**

**Materials are openly  
licensed: CC-BY**



Photo by [Anthony Da Cruz](#) on [Unsplash](#)

# Where does this session fit within the course?



## Course (5 parts)

**Session 1 - What is FAIR RDM and why should we do it? (*Beginner*)**

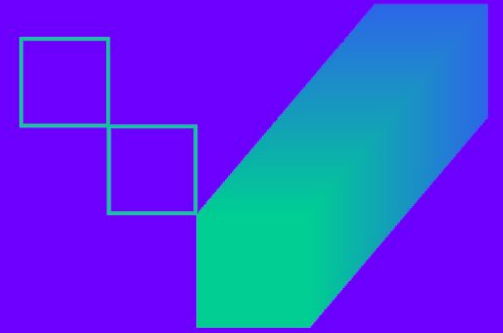
Session 2 - Planning for FAIR: Introduction to RDM and DMPs (*Beginner*) 

Session 3 - Getting started with putting FAIR RDM into practice (*Beginner*)

Session 4 - A deeper dive into putting FAIR RDM into practice - 1 (*Intermediate*)

Session 5 - A deeper dive into putting FAIR RDM into practice - 2 (*Intermediate*)

# Session 2 - topics



- Why RDM and who should do what?
- What is a DMP and how do you create one?
- The research data lifecycle and when to create a DMP
- Funder requirements
- Content of a good DMP



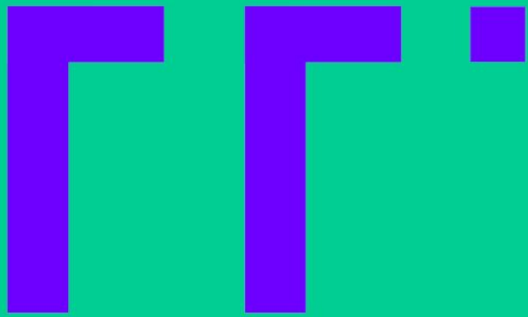
Funded by  
the European Union

# Learner introductions



**Our open slide show (a selection...more next week)!**



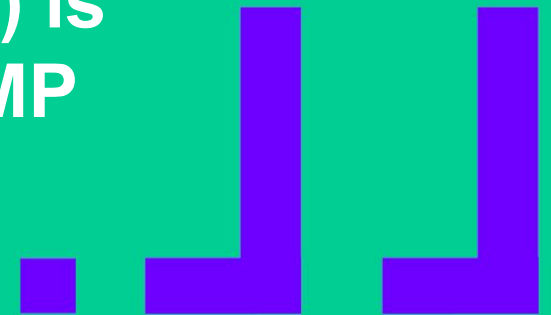


# Learning Outcomes:

For 'Planning for FAIR: Introduction to RDM and DMPs'

By the end of Session 2, you will be able to:

- Define **research data management (RDM)** and **describe its relevance and benefits**
- Explain **why RDM is a step towards FAIR**
- Describe **what a Data Management Plan (DMP) is**
- Paraphrase the **general areas covered by a DMP**



# Course glossary

If a term comes up  
that you would like  
added we can add it!

*[Link deleted: reusers of this  
content may choose to create  
their own course glossary to  
keep updated]*



# Opportunity to Reflect

**How are you progressing with your project work**

Has it brought up any questions or issues?



# PATTERN.

Empowering Open and Responsible  
Research and Innovation

## Why RDM and who should do what?

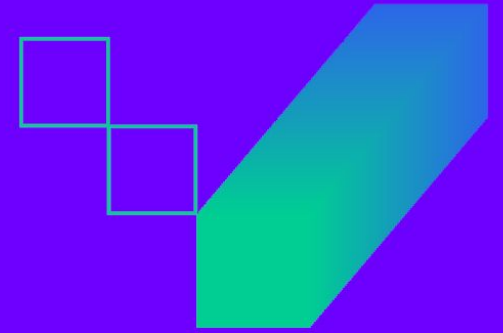
Part of: Session 2: Planning for FAIR: Introduction to RDM and DMPs (Beginner Level)

PATTERN



Funded by  
the European Union

# Session 2 - topics



- Why RDM and who should do what? ←
- What is a DMP and how do you create one?
- The research data lifecycle and when to create a DMP
- Funder requirements
- Content of a good DMP



Funded by  
the European Union

# What is Research Data Management (RDM)?

Handling research data effectively and appropriately throughout the life of a research project and beyond

Research data management refers to all aspects of **creating, storing, sharing and archiving data** and is an **essential aspect of conducting responsible research**

Source: 'Data Management Planning', MANTRA,  
<https://mantra.ed.ac.uk/datamanagementplanning/>

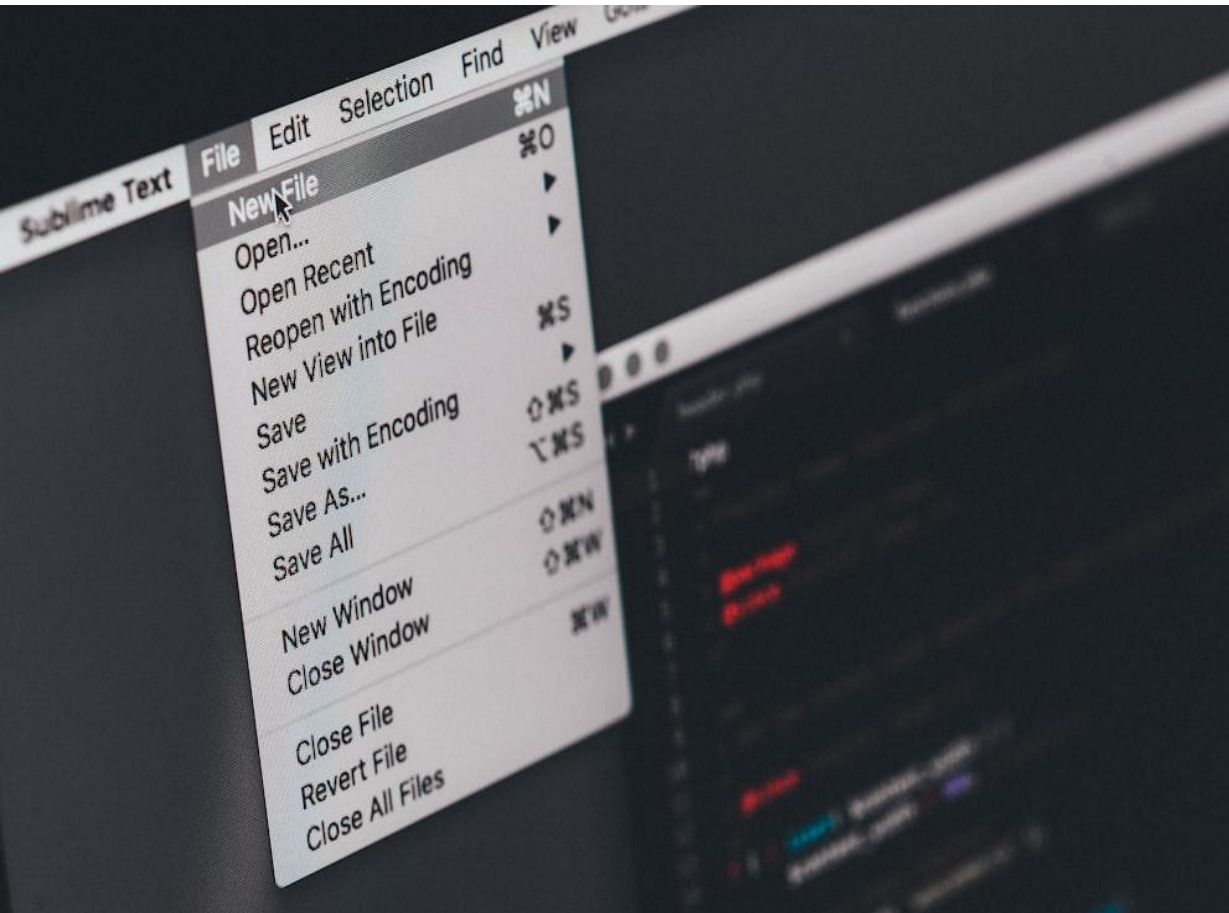


Photo by [Jessica Lee](#) on [Unsplash](#)





# Goals and motivations for managing your data



Have you ever lost data, e.g. through the accidental loss or damage of hardware or accidental deletion or saving over data?

--> data loss or corruption is a big risk. Managing your data well can help avoid this!

# Goals and motivations for managing your data



- If you found a computer folder or physical notebook from 2 years ago – **would you be able to understand the data?**
- If you tried to access computer files that were created 5 years ago, **would you be able to open them?**
- If you published a research article and another scholar contacted you to ask for access to the data underlying your findings – **would you be able to / feel confident about sharing your data with them?**

--> if you answered no to the three questions above, it's time to look into your data management practices!

# What is a Data Management Plan?



A DMP is a formal document that provides a framework for how to handle the data material during and after the research project.





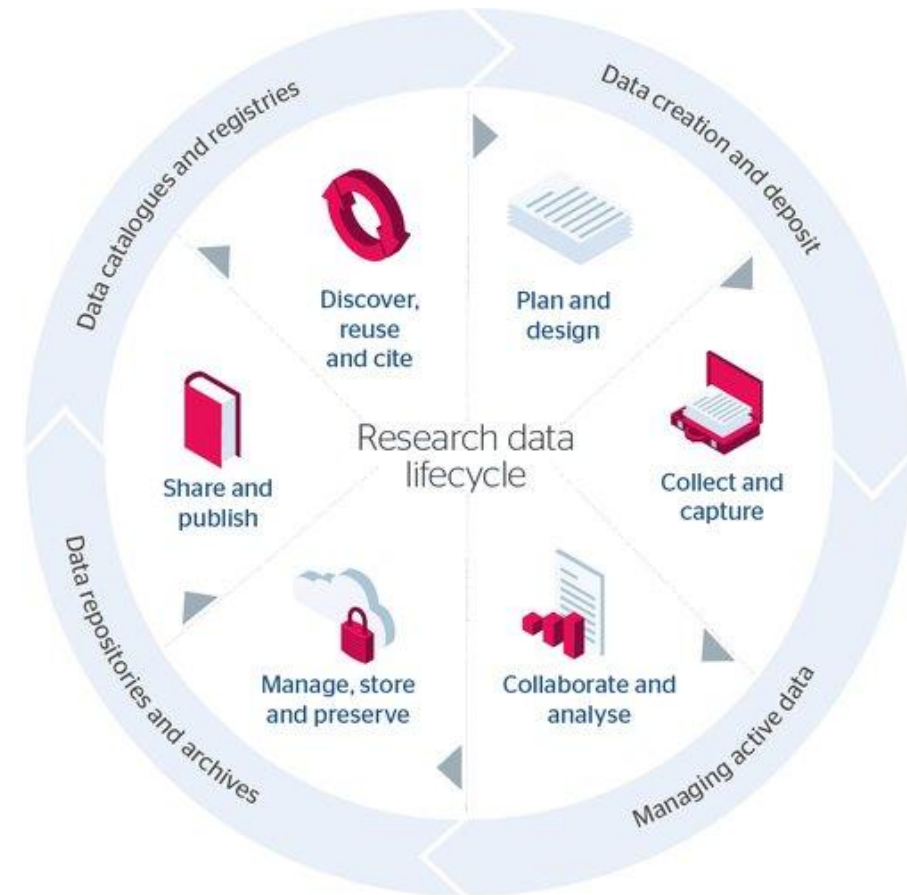
# Who does what in Research Data Management?



Who will create,  
implement and update the  
Data Management Plan?

- Make this **a priority**
- Make it **a collaboration**
- Make a plan for **reviews and updates**

Make sure whoever is responsible for the DMP  
can ensure that it's implemented



# Who does what in Research Data Management?



Who will do each data management activity?

--> From the outset, you need to define roles for each task

Do you need any expertise that is not available in your team?

--> who do you contact for advice, are there any skills that can/should be budgeted for?

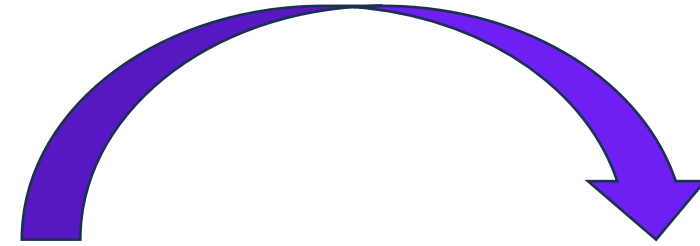


Photo by [Kelly Sikkema](#) on [Unsplash](#)

# Assigning responsibilities



Steps for assigning data management responsibilities:



- For each task identified **in your data management plan**, identify the skills needed to perform the task
- Match skills needed to available staff and identify gaps
- Develop training/hiring plan
- Develop staffing/training budget and incorporate into project budget
- Assign responsible parties and monitor results

You need a data  
management plan!



# Institutional services and support with RDM

**Training** is often available on the general principles of RDM, the FAIR Principles, as well as more discipline-specific approaches and tools

**Support** is often available with planning RDM and creating Data Management Plans

Consider the time, skills, and support you may need with **creating metadata** – a time consuming but important process

*--- take advantage of these supports and training as soon as possible!*



# Don't go it alone!

'Few people have all of the skills required to manage and share data throughout its lifecycle, so **seek input from others** with relevant expertise and **use tools provided by your community**. Don't go it alone —'

Source: 'How to develop a Data Management Plan',  
DCC, <http://www.dcc.ac.uk/sites/default/files/documents/publications/reports/guides/How%20to%20Develop.pdf>

Photo by [Matteo Vistocco](#) on [Unsplash](#)





# Advice and support

--> we will go to the Particify

*[Question we used on Particify: Who do you go to for advice and support on data management and data management plans?]*

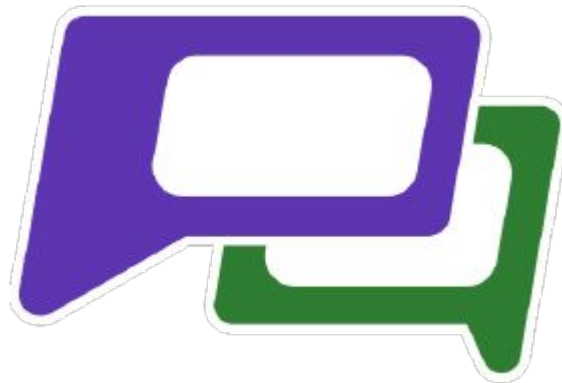


Photo by [Jon Tyson](#) on [Unsplash](#)

## Why Particify?

The basic functionality of Particify is free and can be used at any time via their [web service](#). This is open source software, which they make available on [Gitlab](#) under the respective licenses.



# Developing a DMP helps to avoid data management nightmares!

Coming next: 'what is a DMP and how to create one...'

Source: 'FOSTER - Managing and Sharing Research Data (v2)', <https://openplato.eu/blocks/catalog/detail.php?id=60>



Funded by  
the European Union





# A video to watch in your own time...



## 'Data Sharing and Management Snafu'

Educational video from 2012 that illustrates the value of open data and data standards, that stands the test of time



### Data Sharing and Management Snafu in 3 Short Acts



NYU Health Sciences  
Library  
913 abonnees

Abonneren

1,1K



Delen



<https://www.youtube.com/watch?v=N2zK3sAtr-4>

# PATTERN.

Empowering Open and Responsible  
Research and Innovation

## What is a DMP and how and when do you create one?

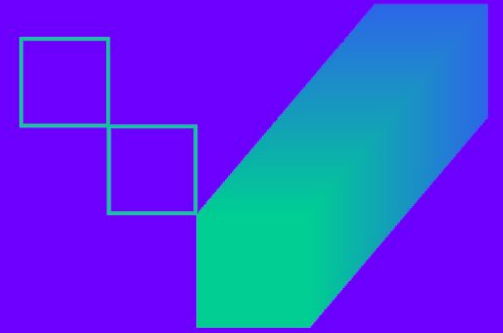
Part of: Session 2: Planning for FAIR: Introduction to RDM and DMPs (Beginner Level)

PATTERN



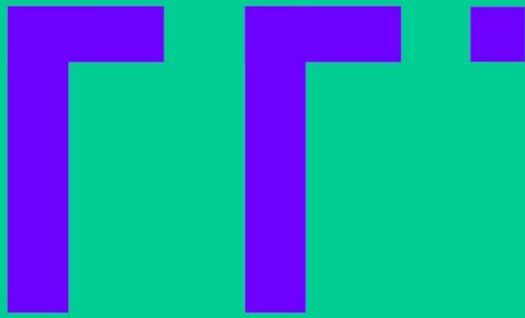
Funded by  
the European Union

# Session 2 - topics



- Why RDM and who should do what?
- What is a DMP and how do you create one? ←
- The research data lifecycle and when to create a DMP ←
- Funder requirements
- Content of a good DMP





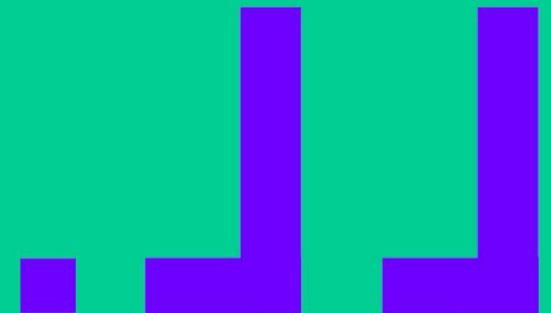
'A goal without a  
plan is just a wish'

[Antoine de Saint-Exupéry  
(1900 -1944)]



Funded by  
the European Union

Source: 'FAIR Data & Data Management Plans', FOSTER,  
<https://www.fosteropenscience.eu/node/2769>





# Why a Data Management Plan?



Information regarding your data management needs to be easily found and understood, not least if you are working on a project that runs over several years and involves a large team of people.



# What is a Data Management Plan?



In order to simplify data management, a Data Management Plan (DMP) can be created early in the research process.

A DMP is a formal document that provides a framework for how to handle the data material during and after the research project.



# What is a Data Management Plan?



The way a DMP will look once it is finished is not universal. It is a **"living" document** that changes together with the needs of a project and its participants.

It is **updated throughout the project** to make sure that it tracks such changes over time and that it reflects the current state of your project.



# How the DMP relates to FAIR



A good data management plan covers all FAIR principles (Findable, Accessible, Interoperable, Reusable)

**A DMP helps to make the data Findable** (F principle) because it includes all information about where data is stored and preserved, during and after the project

**A DMP helps to make the data Accessible** (A principle) because it also includes information about how data can be accessed, what is required to access the data (authentication or authorisation) and by what (standardised and universal) communications protocol, e.g. HTTP, HTTPS

Source: Engelhardt, C. (2022). How to be FAIR with your data.  
<https://doi.org/10.17875/gup2022-1915>, lesson plan 2



# How the DMP relates to FAIR

**A DMP helps to make the data Interoperable** (I principle), because it guides you to think about, e.g.:

- Metadata standards to describe and structure data
- Use of vocabularies in metadata to represent the content, and that these themselves follow the FAIR principles, so that others, humans or machines, can find, access, interoperate and reuse them

'The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.'

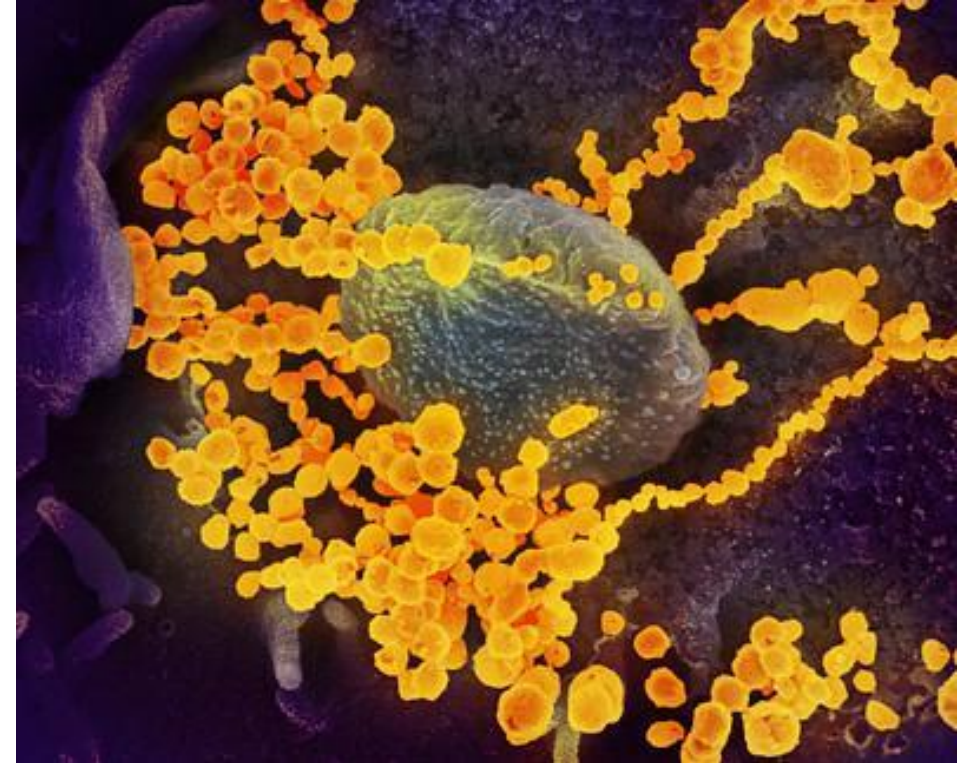


# How the DMP relates to FAIR

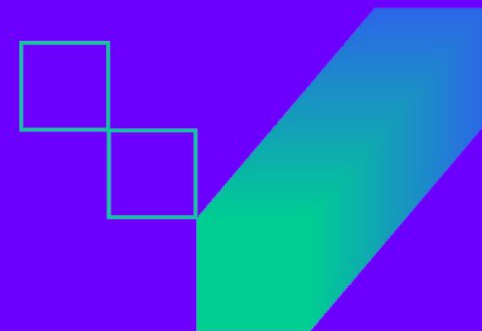


**A DMP helps to make the data Reusable** (R principle) because it allows data to be described with more detail and accuracy, making it easier for others to understand.

Moreover, during DMP creation, it is necessary to indicate the information that is needed to prepare the data for sharing and reuse **with appropriate licences and rules**, namely, how the data can be reused and for whom the data may be valuable



“Novel Coronavirus SARS-CoV-2” by NIAID,  
here cropped, is licensed CC BY 2.0.

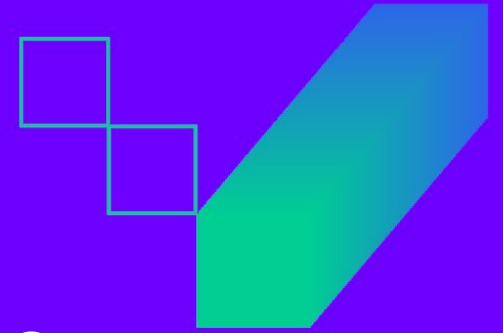


# Content of a good DMP



Funded by  
the European Union

# Template



Check if a specific template is available to you / required of you from particular institutions or funders

Most templates cover the following areas:



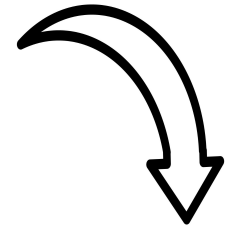


# Description of data

- Will you **generate new data**?
- What **existing data sets** will you **reuse**?
- What **type** of data, **format**, estimated **size**?

--> you can create a table

Let's look at an example!



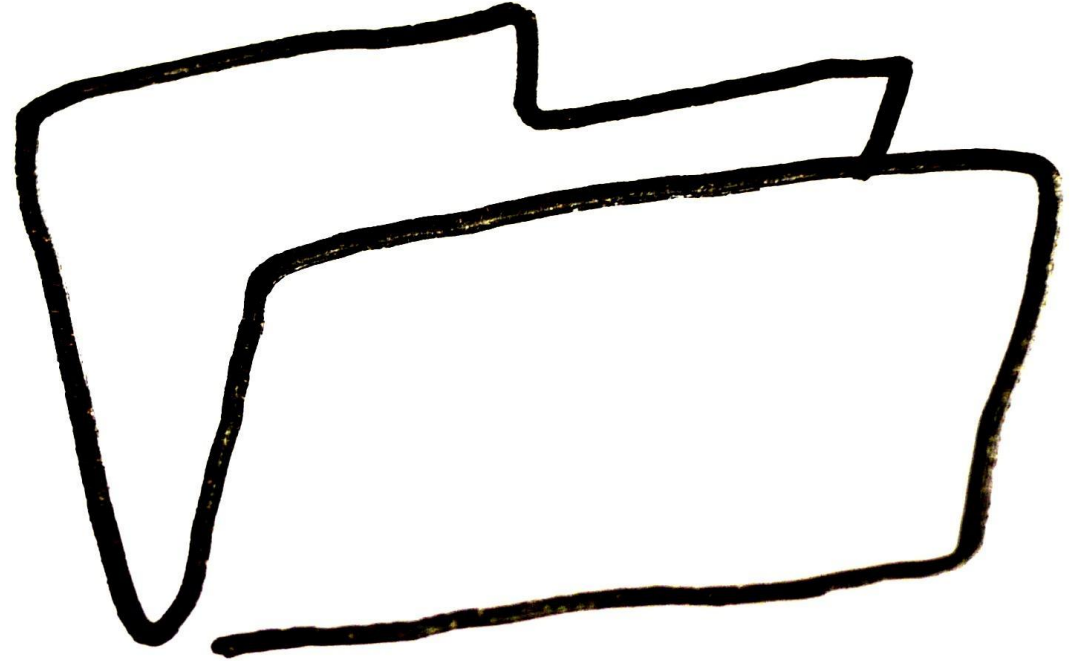
<https://www.kuleuven.be/rm/en/guidance/data-standards/example-data-tables>

# Organisation of the data during the project and in datasets



- File naming conventions
- Versioning
- Folder structure

--> Know what your file is before you double click!



Name	
Press release for approval.doc	
Press release final.doc	
Press release FINAL VERSION.doc	
Press release FINAL FINAL VERSION.doc	
IMPROVED FINAL PRESS RELEASE.doc	
REVISED APPROVED FINAL PRESS RELEASE.doc	
REVISED APPROVED FINAL PRESS RELEASE v. 2.doc	
!! NEW REVISED APPROVED FINAL PRESS RELEASE v. 2.doc	
!!! REVISED NEW REVISED APPROVED FINAL PRESS RELEASE v. 2.doc	
!!!! Press release as sent.doc	

<https://properdiscord.com/2012/09/13/never-use-the-word-final-in-a-file-name/>

# Metadata and metadata standards, Documentation



- What information is needed to make data **understandable** and **findable**?
- Will this be provided using **agreed standards**?

Let's park this here - metadata and documentation is covered in more detail in Session 3 of FAIR Research Data Management!

# Data quality procedures during data collection, data processing, data sharing and reuse



How the consistency and quality of data collection will be controlled and documented. Could include:

- calibration of devices:

*e.g. 'The CO2 sensors will be cross calibrated against atmospheric pressure'*

- repeated samples or measurements

*e.g. 'Data consistency will be assessed by comparing repeated measures.'*



# Data quality procedures during data collection, data processing, data sharing and reuse



- Data entry validation

<https://support.microsoft.com/en-us/office/apply-data-validation-to-cells-29fecbcc-d1b9-42c1-9d76-eff3ce5f7249>

- Peer review of data

- Tidy Data principles

e.g. Tidy Data for Librarians:

<https://librarycarpentry.org/lc-spreadsheets/>

Budget Input -- Human Resources				Budget Maximum	65,000
Cost Center	Category	Current Year Projected	Future Year Budget	Variance	
110	Salary - Regular Wages	63,578	60,000	(3,578)	
abc	Salary - Hourly Wages	19,721	25,000	5,279	
130	Salary - Commission	37,500	40,000	2,500	
140	Invalid Input	X		0	
150		0		1,609	
160		0		1,197	
170		0		5,015	
180		0		713	
190	Tax - SDI	1,470	1,500	30	

# Ethics and intellectual property



- **Ethical approval** when working with personal or otherwise confidential data, or other reasons to seek ethical approval
- Working with **third party data**
- Planning to seek **patent protection for your data**

# Data storage and backup

Where will the data be stored?

How will you ensure that data is protected against risks?

E.g. data loss due to human errors, computer failures or malicious attacks. And how to make sure that it cannot be accessed, deleted or modified by unauthorized persons?

FAIR Data Austria (2021). “Data Management Plan (DMP)”. In: Research Data Management Open Educational Resources Collection. (<https://fair-office.at/index.php/dmp/?lang=en>).



# Data sharing (data access and reuse)



What about preservation and sharing?

- What data will be preserved after the end of the project?
- For how long and where?
- Will the data shared open access? Or there are restrictions to access, what are they?
- What licenses will be applied?

- See: 'How can I use this data? The importance of licences to facilitate reuse', <https://doi.org/10.5281/zenodo.10986072>





# Coming back to 'reasons why not to share data' from Session 1



- Are the reasons we came up with **valid reasons not to share the data at all?**
- Is there **a way that we could plan to share the data?** If so, what measures need to be taken?
- If the data really cannot/should not be shared, **what about the metadata?**

Examples:

- Concerns about anonymisation / re-identification risk
- Preparing and sharing data is time consuming
- Data poorly organised and too much work to make reusable
- Consent to share was not given by participants
- Lack of knowledge

# Responsibilities for managing data and resources

- Who has overall responsibility for data management?
- Who does specific tasks such as creating documentation, preserving and sharing data?
- What training is required?
- Are there any costs involved in data management and sharing?



# Discussion...



What do you think is the most difficult aspect of data management to plan?

Is this also the same aspect to actually *execute*?



# The Research Data Lifecycle and when to create a DMP



- Draft a first version of the DMP to outline your strategy *(as early as possible)*
- Update the DMP along your project with more details.
- Use the DMP as guidelines for data management in your project.
- Discuss and review the DMP regularly with collaborators.
- Keep the DMP along with other project documentation.



## Further steps with your DMP



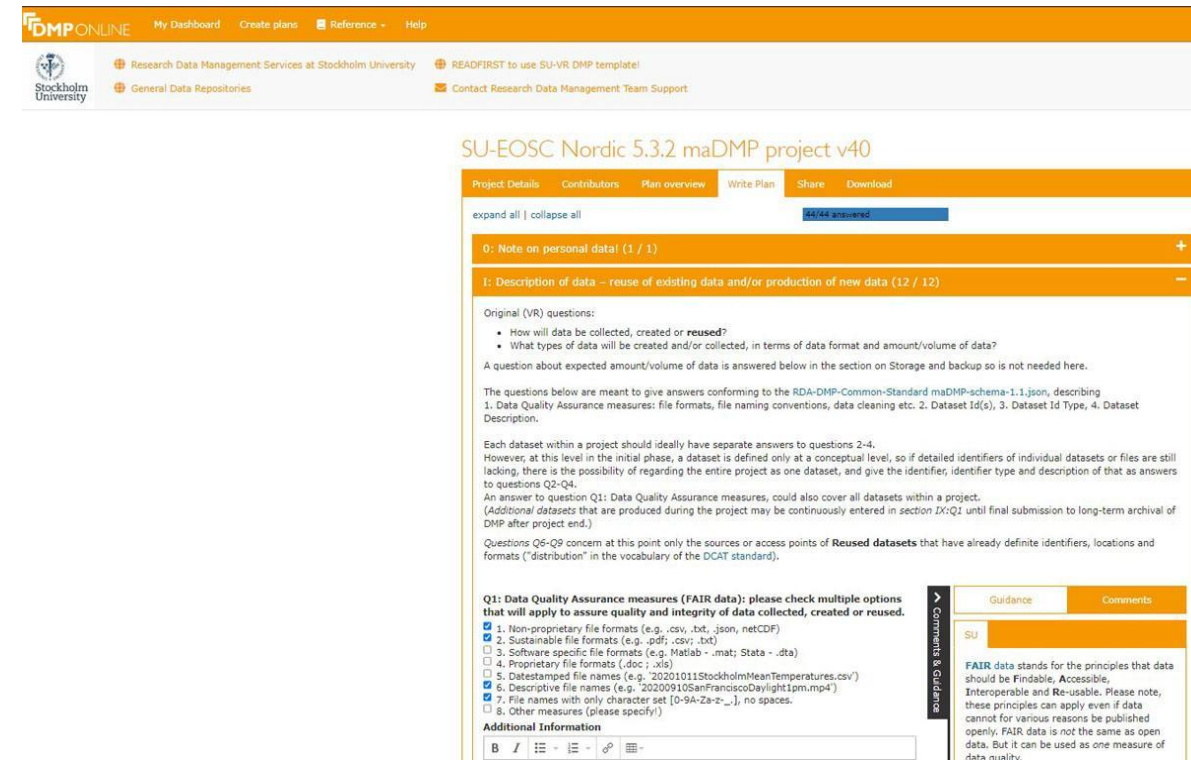
- DMPs can be published (research output)
- DMPs themselves can be made FAIR:

Get a DOI for your DMP by publishing it in Zenodo, etc

# Make your DMP machine actionable

See:

Making Data Management Plans Machine Actionable: Templates and Tools, <https://doi.org/10.5334/dsj-2023-029>, figure 1.



Miksa, T, Simms, S, Mietchen, D and Jones, S. 2019. Ten Principles for Machine-Actionable Data Management Plans. *PLoS Comput Biol*, 15(3): e1006750. [Available at: DOI: <https://doi.org/10.1371/journal.pcbi.1006750>]



# Questions



Funded by  
the European Union







**PATTERN.**

Empowering Open and Responsible  
Research and Innovation

# Session Close

## Session 2: Planning for FAIR: Introduction to RDM and DMPs ?

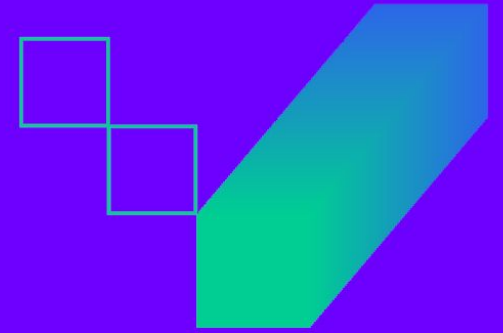
(Beginner Level)

Course: FAIR Research Data Management  
**PATTERN**



Funded by  
the European Union

# Taking your learning a step forward...



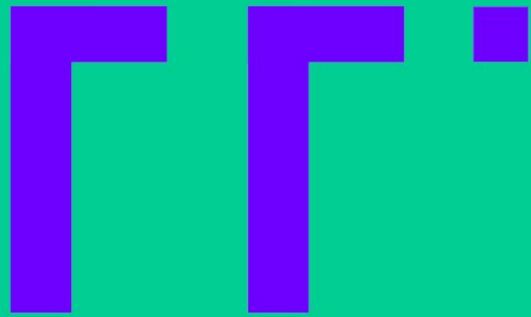
*In your own time...*

Go back to the resource where you just did the quiz and watch the video on 'What is the Content of a DMP?' (Duration: 5:23 mins)

URL: [https://fair-office.at/?page\\_id=5761&lang=en](https://fair-office.at/?page_id=5761&lang=en)



Funded by  
the European Union



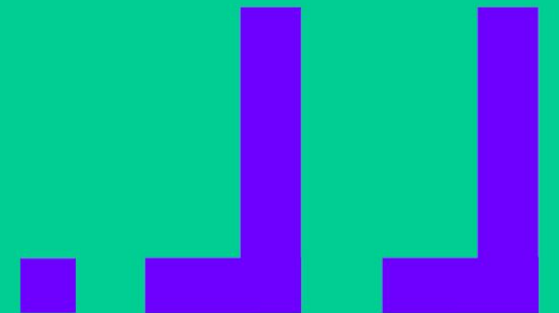
# Coming next:

The next session in the 'FAIR Research Data Management' Course is:

## 'Session 3: Getting started with putting FAIR RDM into practice'

Learn about:

- Metadata and data documentation
- The basics of file formats
- How to publish and share data



# Take home messages

--> we will go to the Particify

*[Question we used on Particify: What is your biggest take-home message?]*



Photo by [The New York Public Library](#) on [Unsplash](#)

# PATTERN.

Thank you!

Deborah Thorpe, Loek Brinkman, Michelle van den Berk, and René van Horik,  
DANS

## OUR CONSORTIUM

