



# Research Data Management for the Life Sciences

Emily Kate and Michael Feichtinger



Cite as Feichtinger, Michael and Emily J. Kate. 2023. University of Vienna Research Data Management for PhD Candidates in the Life Sciences Course Workshop Slides. Version 01. [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/).

# Workshop Schedule

## Day 1

9:00-9:30 Welcome and Introduction  
9:30-10:20 Research Data Life Cycle, FAIR data principles, Open Research  
10:20-10:35 Break  
10:35-11:10 Data Management Plans  
11:10-11:45 Organizing your Data  
11:45-12:00 Storing and Securing your Data (Part I)

## Day 2

9:00-9:15 Recap and Open Questions of Day 1  
9:15-9:30 Storing and Securing your Data (Part II)  
9:30-10:00 Documenting and Annotating your Data  
10:00-10:10 Break  
10:10-11:10 Publishing and Archiving your Data  
11:10-11:15 Break  
11:15-11:45 Revisit DMP and Course Evaluation  
11:45-12:00 Research Ethics and Legal Issues



# Learning Objectives Day 1

- Explain the research data life cycle, FAIR principles and Open Research
- Write a data management plan for your project
- Develop a data organization schema for a research project
- Know how and where to securely store your data



Get to Know

Assign each group member a circle and write their name on the corresponding line.

Write down two professional facts that make each person unique.

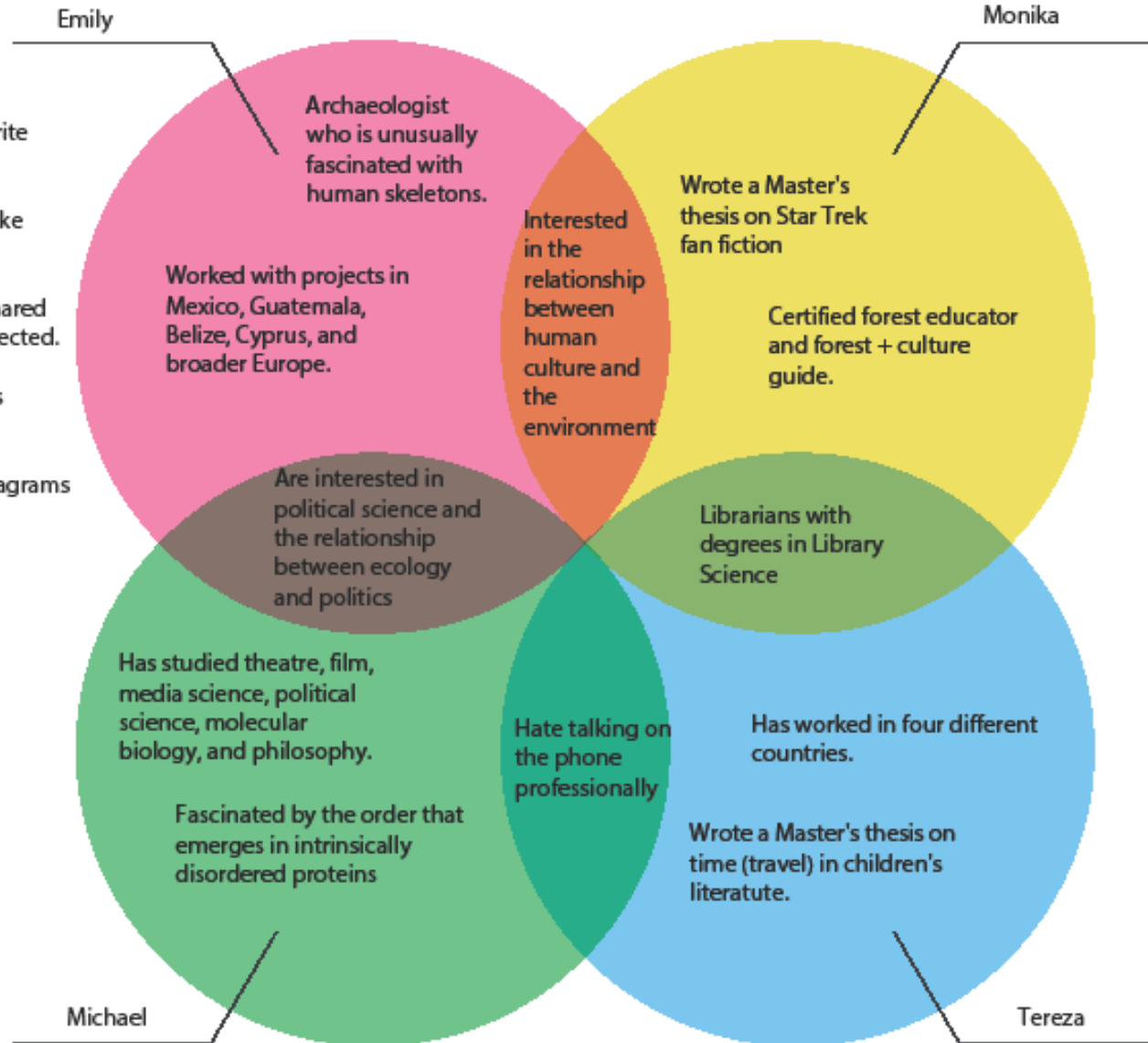
Write down one professional fact that is shared between members whose circles are connected.

If you can, find one professional fact that is shared by all members of the group!

We will present the results of your venn diagrams once everyone is done.

Our shared fact:

We all work as research data management specialists!





# What is Research Data Management?

Research Data Lifecycle, Fair Data  
Principles, and Open Research



# What is Research Data?

**“Research data** comprise all materials that are created in the course of academic work, including records, source research, experiments, measurements, surveys and interviews, as well as software and code. Research data may take on various different forms: during the lifespan of a research activity, data may exist as gradations from raw data to processed data (and even include negative and inconclusive results).”

- RDM Policy of the University of Vienna



# What is Metadata?

- Data about your data
- Different types of metadata
  - Administrative metadata
  - Structural metadata
  - Descriptive metadata



# Activity: Research Data Life Cycle

## **Group work:**

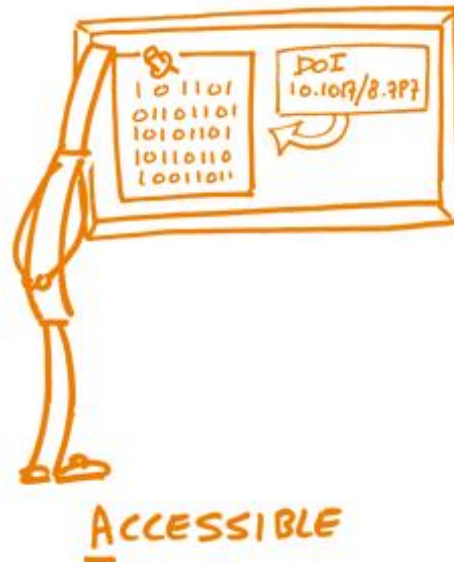
Arrange the research data life cycle cards in the order that makes most sense to you.

Add steps (blank cards) if you think something is missing.

Allocated time: 5-10 min

# FAIR Data Principles

## FAIR DATA PRINCIPLES





# Open Research

Open research is the practice of making “the primary outputs of publicly funded research results – publications and the research data – publicly accessible in a digital format with no or minimal restriction.”

- Organisation for Economic Co-operation and Development (OECD)

# Activity: FAIR and Open Research Cost Benefit Analysis

## **Group work:**

Discuss the assigned question(s) regarding FAIR data and/or Open Research in your group.

Allocated time: 10-15 min



# Data Management Plans

# What is a DMP?

A DMP “describes how research data should be handled, and is therefore a tool for research data management. The plan explains which data are collected or generated during a research activity and specifies what needs to be done during the data lifecycle, i.e. the whole data process (storage, publication, citation, long-term availability, pseudonymisation, anonymisation, reuse, deletion, etc.)”.

- RDM Policy of the University of Vienna

# Activity: DMP Journal Time

Try to add information or notes on your own research project to each section of the DMP.

Allocated time: 15 min




# Organizing your Data





# Data Organization

- Manage your raw data
  - Use effective file names
  - Design a clear and consistent folder structure
  - Select proper file formats
- 

# Activity: Data Organization Schema for a Research Project

## **Group work:**

Read your project scenario and develop a folder structure that represents the different data types in the project.

Come up with file naming conventions that contain the crucial information for the data types.

*Don't worry if you are not familiar with the experimental procedures in the projects! Focus more on how to organize data from different sources and/or temporally or spatially distributed series of experiments.*


Allocated time: 20 min



# Storing and Securing your Data



# Best-Practices and Resources

- 3-2-1 concept for back-ups
  - u:cloud or OneDrive
  - Institutional shares
  - Back-ups on LiSC
  - Sensitive Data
- 



# Questions? Concerns? Comments?



# Research Data Management for the Life Sciences

Emily Kate and Michael Feichtinger

# Workshop Schedule

## Day 1

9:00-9:30 Welcome and Introduction

9:30-10:20 Research Data Life Cycle, FAIR data principles, Open Research

10:20-10:35 Break

10:35-11:10 Data Management Plans

11:10-11:45 Organizing your Data

11:45-12:00 Storing and Securing your Data (Part I)

## Day 2

9:00-9:15 Recap and Open Questions of Day 1

9:15-9:30 Storing and Securing your Data

9:30-10:00 Documenting and Annotating your Data

10:00-10:10 Break

10:10-11:10 Publishing and Archiving your Data


11:10-11:15 Break

11:15-11:45 Revisit DMP and Course Evaluation

11:45-12:00 Research Ethics and Legal Issues



# Learning Objectives Day 2

- Know how and where to securely store your data
  - Write a readme for your project and data
  - Find a proper repository to publish and archive your data
  - Choose licenses for your published data and code
  - Be aware of ethical and legal issues that you have to consider in your data management
- 






Any open questions from yesterday?



# Storing and Securing your Data



# Best-Practices and Resources

- 3-2-1 concept for back-ups
  - OneDrive or u:cloud
  - Institutional shares
  - Back-ups on LiSC
  - Sensitive Data
- 

# Activity: Storage and Security Decisions

## **Group work:**

Sort the handed-out decisions regarding back-ups, data security and the handling of sensitive data from bad to good, better, and best.


Allocated time: 10 min



# Documenting and Annotating your Data



# How and Why Document your Data?

- Someone who is familiar with your field can interpret your data.
  - Project- and data-level documentation
  - Metadata standards
  - README files
- 

# Activity: README Journaling


Try to add information or notes on your own research project or one of your data sets to each (relevant) section of the README template.

Add additional sections or information that may be relevant to understand your project or data set.

Allocated time: 10-15 min



# Version Control

- ... is a system that **keeps track of changes** made to files, allowing users to **easily revert** to previous versions, **collaborate** on projects, and manage different versions of the same files.
  - **Basic**: Informatively named files
  - **Automated**: Built in Software Platforms
  - **Advanced**: Version Control System such as Git
- 






# Publishing and Archiving your Data



# Why Publish and Preserve your Data?

- **Verification** and **reproducibility** of your data for several years after the end of the project
  - Allow the **reuse** of your data
  - Funders, publishers, institutions and organizations requirements
  - Significant value for society
- 



# How to Publish and Preserve your Data?

- Choose a repository
- Get a persistent identifier
- Assign a license

# Example: Finding Repositories

Your research project aims to understand the mutualistic relationship between lucinid clams and sulfur-oxidizing bacteria in their gill cells. Therefore, you will employ a combination of **microscopic observation**, including spatial arrangement of the bacteria, and histopathological parameters, with molecular techniques, including **qPCR**, and **metagenomic sequencing**, to elucidate the mechanisms of this host-microbe association.


















# Activity: Finding a Repository for your Data

## **Group work:**

Try to find suitable repositories for the data of each person in your group.

Allocated time: 15 min

# Licensing your Data

CREATIVE COMMONS LICENSES		 COPY & PUBLISH	 ATTRIBUTION REQUIRED	 COMMERCIAL USE	 MODIFY & ADAPT	 CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✓
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✓
 You can redistribute (copy, publish, display, communicate, etc.)		 You have to attribute the original work		 You can use the work commercially	 You can modify and adapt the original work	 You can choose license type for your adaptations of the work.

# Licensing your Data

License

required ▼

Access right \*

☐ Open Access

☒ Embargoed Access

☐ Restricted Access

☐ Closed Access

Required. Open access uploads have considerably higher visibility on Zenodo.

📅 Embargo date

Required only for Embargoed Access uploads. Format: YYYY-MM-DD. The date your upload will be made publicly available in case it is under an embargo period from your publisher.

⚙ License \*

Start typing a license name...

Creative Commons Attribution 4.0 International

Creative Commons Attribution 1.0 Generic

Creative Commons Attribution 2.0 Generic

Creative Commons Attribution 3.0 Unported

10.11.10.2023

RDM for the Life Sciences

39

# Activity: Licensing Scenarios

## **Group work:**

Try find solutions to your assigned research scenarios that include details about licensing requirements, software, and data reuse.

Allocated time: 10 min



# Course Evaluation

Complete the post-learning assesment and the course evaluation on Moodle.

Put your name and e-mail into the shared excel sheet on Moodle to receive a confirmation for your participation.

Allocated time: 15-20 min

# Activity: Revisit your DMP

Revisit your DMP draft from yesterday and review or add further information based on the knowledge you obtained in the last two days.

Allocated time: 10 min

If you would like to receive feedback on your DMP, please either schedule a meeting with one of us or upload your completed DMP to Moodle (till October 31).




# Ethical and Legal Aspects of RDM

# Research Ethics and Research Integrity

- **Responsible Research and Innovation** is defined as “an approach that anticipates and assesses potential implications and societal expectations with regard to research and innovation, with the aim to foster the design of inclusive and sustainable research and innovation.” (European Commission)
- **Research Integrity** refers to an “active adherence to the ethical principles and professional standards essential for the responsible practice of research” (Korenmann 2006).



# Legal Considerations

- CARE principles
  - General data protection regulation (GDPR)
  - Working with animals
  - Nagoya protocol
  - International transport of samples
  - Research using private funds
- 



# Questions? Concerns? Comments?



Let's keep in  
touch!

[emily.jean.kate@univie.ac.at](mailto:emily.jean.kate@univie.ac.at)  
[michael.feichtinger@univie.ac.at](mailto:michael.feichtinger@univie.ac.at)  
[rdm@univie.ac.at](mailto:rdm@univie.ac.at)