**Data management planning Episode Notebook**

Part of FAIR in (bio) practice, <https://carpentries-incubator.github.io/fair-bio-practice>

**Attendance**

Please sign in here: your name/ pronoun (if you prefer to share it) / institution / country.

1.

2.

3.

4.

5.

6.

7.

8.

**Exercise 1: Action plan challenge**

Where would you say the following actions belong in the Research Data Life Cycle? How do they help in achieving FAIR principles? (type the corresponding number of the research data Life Cycle after each corresponding action)

Research Data Life Cycle

* Creating data
* Reusing data
* Processing data
* Analysing data
* Preserving data
* Sharing data

Actions:

* Clarify usage rights:
* Give credit through citations:
* Use open source software:
* Attach PID to your data:
* Attach descriptive metadata:
* Produce standard metadata:
* Backup your data:
* Create figures and plots in python/R:
* Organize your files in folders:
* Select data repository:
* Add open licence:
* Link publications, data and methods:
* Create a template for assay description:
* Use institutional repositories:
* Use controlled vocabularies:
* Convert numerical data to csv:
* Track versions of files:
* Performing statistical analysis:
* Deposit datasets to Zenodo/Dryad:
* Record experiment details in Electronic Lab Notebook:
* Use github for your code:
* Ask someone to revise your project structure:
* Reformat and clean data tables:
* Use a Minimal Information Standard:
* Use PID in data description:
* Download a dataset:
* Link to UniProt or GenBank records:

**Exercise 2: Challenge**

Working in groups, think of your last paper (or project). Pretend that you have a joined project that combines the outputs of at least two your papers/projects.

You can look at the example of DMP and resuable paragraphs:

<https://www.wiki.ed.ac.uk/display/RDMS/Short+paragraphs+that+you+might+find+useful+when+preparing+your+DMP>

Our list of suggested Data Repositories can be found here: <https://www.wiki.ed.ac.uk/display/RDMS/Suggested+data+repositories>

For finding standards and repositories:

<https://fairsharing.org/>

For ontologies: <http://www.obofoundry.org/> , <https://bioportal.bioontology.org/>

Write a short DMP for this **joined project**.

**Drop the DMP document at:** <https://uoe-my.sharepoint.com/:f:/g/personal/tzielins_ed_ac_uk/EknZgNnsY-tOjWKtewc1oGcBfg4cyb-R65lQvANqkR5nog?e=1If3ZL>

HINT: You can drop a document there and start to collaborate on it online doing simultanous edits!

Your DMP should contain the following three sections:

1. What data you will acquire during the project: Please describe the type of data you will generate (for example ‘flow cytometry data’) as well as file formats and data volume. these data will be stored under (include the meta data as well). Estimate the size of your data.

-

-

-

1. How you will store the data: Please describe how you will store and organize your data, what metadata will you capture in a what form. Tell how you or document the data during the duration of the project

-

-

-

1. How you will share the data: Please describe the strategies for data sharing, licensing and access information.

-

-

-

**Remember: it is a joined project**

**Exercise 2, part 2: check out each others DMPs**

Now, take a look at the other group's DMP and make comments/suggestions on how to improve it (at the end of the DMP)

Green Room <-> Blue Room

Red Room <-> Yellow Room

**Exercise 3: Data management planning Quiz**

Which of the following statements about data management and planning are true/false? (type T or F next to each statement)

1. The best time to do data management is at the end of a project, when you've collected all the data you're managing.
2. Data management plans (DMPs) detail what will happen to data before collection begins.
3. The best storage method for data is multiple backups to USBs.
4. There is a single best way to manage, organise, and share data.
5. For grant applications, DMPs should mention data preservation, longevity, sharing, discover, and reuse.
6. Your metadata should be standardised and descriptive.
7. Taking the time to plan out what's needed in metadata and your DMP will save you time in the long run and make your data more FAIR.
8. DMP online is a tool which constructs DMPs for researchers.
9. Data addressed in a DMP can be freely shared regardless of confidentiality.
10. Data can be given creative commons licenses to dictate how others can and cannot use it.

**Feedback: Help us improve!**

1.      How do you feel about the presented topics after this session (type +1 next to the statement that best describes your feeling):

•       I am more confused:

•       I have a better understanding of them now:

•       My knowledge has not changed much:

2.      Thinking of your knowledge of the lesson topic and its presentation,

which one of the statements best characterize your experience (type +1

next to the statement)

•       I am a novice, and I found the course useful/informative:

•       I am a novice, but I think the course should be improved:

•       I have experience in the presented area, but I found the course

useful/informative:

•       I have experience in the presented area, and I think the course could

be improved:

3.      How was the pace of the lesson:

•       Too fast:

•       About right:

•       Too slow:

4.      If the lesson had to be 5 minutes shorter, what would you remove:

-

-

-

5. If the lesson could be 5 minutes longer, what would you add or spend

more time on:

-

-

-