Data Management

Peter Crowther - 21/07/20

Why data management?

- We collect a lot of data

- Data looses value if not cared for

- Want to achieve open science

- Because we said we would – this is science just like the materials

What is open science?

- Science is not a competition – we all learn faster if we share

- We share inside the group already so why not share between group and between institutions?

- We are already doing high quality work, open science is a way to validate that

Context of a PhD

- You as a PhD student are not responsible for the success of your supervisors grant

- You need to try really hard and act as directed

- These are guidelines – ultimately it's up to you

Data management guidelines

- New data management section on LightForm wiki:

https://lightform-group.github.io/wiki/data_analysis/

- Still working on it – more details soon

- Ongoing projects like Datalight

Organisation of Research Data

- Folders have specific purposes – use them

- Example of Sumeet's data

Metadata

- Metadata should be recorded along with the data

- Use a lab book

- Metadata templates

Uploading data to Zenodo

- Aim is to upload data and analysis to an online repository to be shared with all

- Don't worry about this but bear it in mind

Reproducible workflows

- This is a slightly different thing to data management but feeds into the open science

- Try and do good code

- Try and do traceable analyses – no spreadsheets!

Writing Papers

- Should be able to take the data from the analysis folders straight into the paper