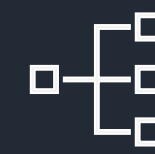


Data management tips



GOAL of good data management
→ optimise the discovery & reuse of data

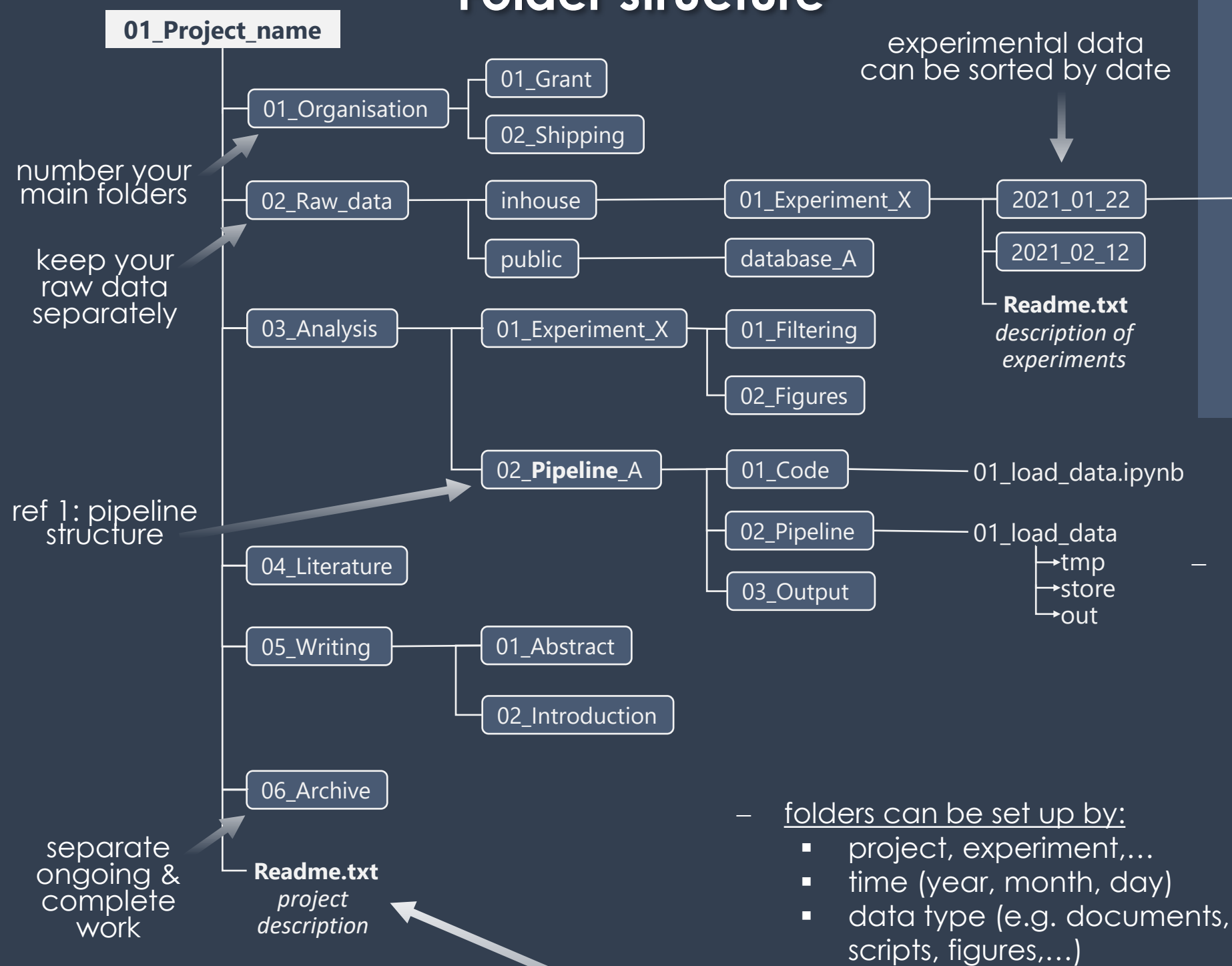


Questions to ask yourself

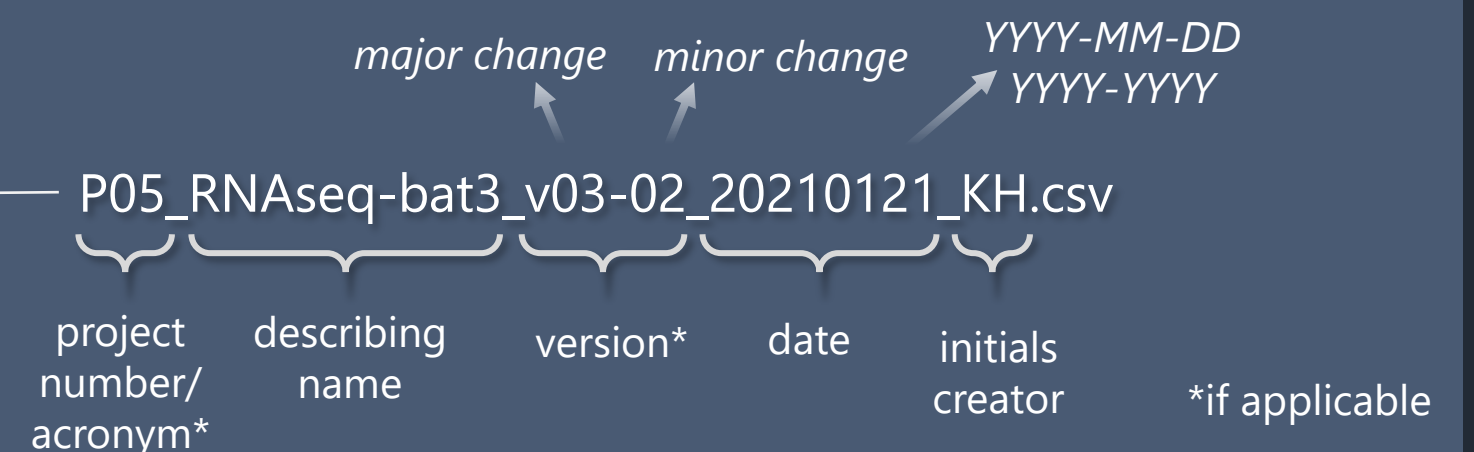
Are my files organised in a way that I can easily find what I am searching for?
What information would I need to understand and use my data in 20 years?
Could others understand and use my data?



Folder structure



File naming



General naming tips for folders & files

- use unique, meaningful names
- not too long (not >30-40 characters)
- no spaces, dots, or special characters (\$"%!'&*"^()+=[:;~@)
- hyphens (-) & underscores (_) to separate elements

- find the balance between a shallow & deep folder hierarchy
 - too deep → too many clicks might be needed to get to the right file
 - too shallow → too many files might end up in one folder (organise them in subfolders)

shallow deep

Friendly Reminder

Comment your code!



Metadata

Which information is necessary to interpret, understand, and use a given dataset?

readme.txt files

can be used to describe projects, folders, and files



Important information

- Who** has created the data?
- What** is the content of the data?
- Which** questions have been answered?
- When** were the data created?
- How** were the data developed (methods)?
- Why** were the data developed?
- With whom** can the data be shared?

References

- ¹ <https://towardsdatascience.com/how-to-keep-your-research-projects-organized-part-1-folder-structure-10bd56034d3a>
- <https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Doing/Organising-files-and-folders.htm>
- <https://www.massey.ac.nz/massey/research/library/library-services/research-services/manage-data/organise.cfm>
- <https://library.bath.ac.uk/research-data/working-with-data/organising-data>
- <https://www.helsinki.fi/en/research/organizing-data-folders-with-5sdata-method>
- <https://mantra.edina.ac.uk>
- <https://old.dataone.org/education-modules>