# Tidy data beyond the dataset: practical tips:

1. Organise your file structures on your computer so that you can easily find what you have saved. Ensure you rename documents so that you can identify them:
   1. Be consistent
   2. Be descriptive
   3. Keep like with like
2. Make sure you back up your research data, documents and associated PDFs. Options include:
   1. External hard drive
   2. Cloud storage e.g. Dropbox or Syncplicity (for confidential data)
   3. Consider security and privacy – how can you keep sensitive data safe?

NOT Pen drive –these can be prone to breaking or being lost

1. Balancing security and access
   1. If you are collaborating with others have a conversation about how you will all manage the data:
      1. Version control
      2. Description/Metadata
      3. Ownership
      4. Long term access
2. Ensure that your files are saved in a format that allows the future you, or a future researcher, to open them:
   1. Where possible use non-proprietary (open) file formats
   2. If you must use proprietary file formats, consider adding a README.txt file with information like the name and version of the software used to generate the file, as well as the company who made the software
   3. Where possible use file formats in common usage by the research community
3. Maintain versioning control so that you can go back and retrieve specific versions of your files later:
   1. Keep a preservation copy of the raw data. Work on a copy.
   2. Consider using a version number in the file name. Example of how to systematically do this: <https://www2.le.ac.uk/services/research-data/documents/UoL_VersionControlChart_d0-1.pdf>
   3. Use built-in options that come with software (but be aware they can increase the file size)
   4. Use version control software e.g. Git
4. Create your own metadata and associate it with your data files. The metadata describes the basic characteristics of the data and makes it easier for you and others to identify and reuse data correctly at a later date:
   1. This can be as simple as adding a README.txt file in the same directory as the data files. Information you could record includes:
   * Title
   * Creator (Principal Investigators)
   * Date Created (also versions)
   * Format (and software required)
   * Subject
   * Unique Identifier
   * Description of the specific data resource
   * Coverage of the data (spatial or temporal)
   * Publishing Organization
   * Type of Resource
   * Rights
   * Funding or Grant
   1. Use the University of Otago Data Management Plan. The DMPtool is an online tool which enables you to store information about your research project and the datasets associated with that project in one place. The DMP tool is free to use, safe, secure and you can access your information anytime from any computer or mobile device: <http://www.otago.ac.nz/library/dmp/index.html>
   2. At an advanced level, depending on your discipline, you may be required to work with a specific metadata schema/standard and use a controlled vocabulary or ontology
5. Plan for your research data long-term:
   1. What data needs to be kept?
   2. Where will you store your finished data once the research is complete?
   3. Do you plan to share your data? Have you accounted for this in your ethics application and consent forms?
   4. Does de-identification need to occur and how will you manage this? This chart provides some useful tips: <https://fpf.org/2016/04/25/a-visual-guide-to-practical-data-de-identification/>
   5. Will you upload your data to a data repository? What would be an appropriate repository for your data?
6. If sharing your data is an end goal, plan how you can help other researchers discover and use it:
7. Can you make it openly available in a data repository?
   1. Does the repository allow you to add a Creative Commons Licence?
   2. Does the repository provide a persistent identifier e.g. DOI?
   3. Does the repository provide support to researchers wanting to know how to cite the dataset?
   4. When completing the metadata record ensure you have words that will help search engines locate your data
8. Use social media to funnel people to your data
9. Get an ORCiD (**https://orcid.org/**) and record your research outputs in your record including datasets.
10. Learn more about FAIR data principles: <https://www.go-fair.org/fair-principles/>
    1. Try at ANDS FAIR self-assessment tool: <https://www.ands-nectar-rds.org.au/fair-tool>

\*\*\*Most of this applies to physical data as well\*\*\*

Want to learn more?

<https://otago.libguides.com/data_management>