# Graduate Student Data Management Plan

**Name:**

**Project Title:**

**Supervisor Name(s):**

**Institution:**

**Date:**

## 1. Data Collection and Organization

### **Types of Data Generated:** List experimental approach and briefly describe data generated and file format. Edit the examples to match your data, add additional data types and delete types you don’t use.

[ ] Real-Time PCR (Ct values, amplification curves in instrument-specific format, exported to .csv, .xlxs, reports in .docx, .pdf)

[ ] Western Blot (Images .tif], [.jpg]) and quantification data e.g., band intensities in [.xlsx])

[ ] Fluorescent Microscopy (e.g., [.tif], [.czi], including metadata like magnification, staining)

[ ] RNA Sequencing (e.g., raw FASTQ files, processed tables in [.csv] or [.xlsx])

[ ] Flow Cytometry (e.g., [.fcs] files, analysis reports in [.pdf] or [.pptx])

[ ] Code (R Studio, [.rmd] with version control through Github. Reports as [PDF]).

[ ] Your approach (describe data types)

### **File Naming Conventions:**

Describe how you will name your data files consistently. Include key information like date, experiment ID, sample ID, and data type. Be specific!

Example: YYYYMMDD\_ExperimentID\_SampleID\_DataType.extension (e.g., 20250506\_Exp01\_CellLineA\_qPCR.xlsx)

List the key elements you will include in your file names, and define any terms/abbreviations:

* + - Date (YYYYMMDD)
    - Experiment Identifier (e.g., Exp01, PilotStudy)
    - Sample Identifier (e.g., Ctrl1, TreatA\_Rep2)
    - Data Type (e.g., qPCR, Flow, Microscopy)
    - Other relevant information: \_\_\_\_\_\_\_\_\_\_\_

### **Folder Structure:**

Describe how you will organize your digital files on your computer/shared drive. Consider how someone would search for a particular piece of data – what information are they most likely to rely on?

Example:

ProjectName/

├── RawData/

│ ├── qPCR/

│ ├── WesternBlot/

│ ├── Microscopy/

│ ├── RNAseq/

│ └── FlowCytometry/

├── ProcessedData/

│ ├── qPCR\_Analyzed/

│ ├── WB\_Quantified/

│ ├── Microscopy\_Processed/

│ ├── RNAseq\_Results/

│ └── FlowCytometry\_Gated/

└── LabNotebook/ (Digital or Physical)

Describe your planned folder structure: \_\_\_\_\_\_\_\_\_\_\_

### **Lab Notebook:**

Will you use a physical notebook, an electronic lab notebook (ELN), or both? How will you link your experimental data to your notebook entries?

[ ] Physical Notebook: How will you reference digital files in your notebook?

[ ] Electronic Lab Notebook (ELN): Which platform will you use? How will you organize your experiments and link data?

[ ] Both: Describe how you will integrate them.

## 2. Metadata and Documentation

### Metadata Standards:

What essential information will you record alongside your data to make it understandable in the future? Think about what someone else (or you in a year!) would need to know to interpret the data. Refer to established guidelines such as MiFlowCyt or MIQE where appropriate.

* For all data types, you will record:
  + Date of experiment
  + Experimenter name
  + Project title
  + Brief description of the experiment's purpose
* For specific data types, you will also record (**edit** to match the approaches/data types/guidelines you will use):
  + qPCR: Instrument used, reagents, primer sequences, cycle conditions, sample preparation details – will follow MIQE guidelines (Bustin et al 2009 - <https://doi.org/10.1373/clinchem.2008.112797>)
  + Western Blot: Antibodies used (primary and secondary, including dilutions and catalog numbers if possible), membrane type, blocking buffer, detection method, loading controls. Will follow Western Blotting Minimal Reporting Standards (e.g. Kroon et al 2022 - <https://pmc.ncbi.nlm.nih.gov/articles/PMC9518894/>, Gilda et al 2015 <https://doi.org/10.1371/journal.pone.0135392>)
  + Microscopy: Microscope model, objective lens, filter sets, staining protocol, software used for image acquisition – will follow QUAREP -Bare Minimal Microscopy Reporting Requirements Checklist (Llopis et al 2025 - <https://zenodo.org/records/14977578>)
  + RNA Sequencing: Library preparation kit, sequencing platform, read length. Will follow MINSEQE guidelines (<https://doi.org/10.5281/zenodo.5706412>)
  + Flow Cytometry: Instrument model, fluorochromes used, antibody clones, gating strategy (briefly) – will follow MIFlowCyt (Lee et al 2008 <https://isac-net.org/page/MIFlowCyt>).
  + Other data types: enter and provide details as in the examples

### Documentation Methods:

How will you ensure metadata is captured and linked to your data files? (check all that apply)

[ ] Separate metadata files (e.g., [.txt] or [.csv] files linked to specific datasets)

[ ] Embedding metadata within file properties (where applicable)

[ ] Detailed descriptions in the lab notebook referencing file names

[ ] Data dictionary (e.g. defines column headers, vocabulary)

[ ] Other: \_\_\_\_\_\_\_\_\_\_\_

## 3. Data Storage and Backup

### Primary Storage Location:

Where will you primarily store your active experimental data? Consider who has access and whether the format will be able to handle the amount of data you are likely to generate.

[ ] Local computer hard drive [provide path]

[ ] Shared lab server [provide path]

[ ] Cloud storage (specify platform: \_\_\_\_\_\_\_\_\_\_)

[ ] CESNET

[ ] External hard drive

### Backup Procedures:

How will you ensure your data is protected against loss or corruption?

*Backup frequency:*

[ ] Daily

[ ] Weekly

[ ] Other: \_\_\_\_\_\_\_\_\_\_\_

*Backup location(s):*

[ ] Shared lab server

[ ] Cloud storage (specify platform: \_\_\_\_\_\_\_\_\_\_)

[ ] External hard drive (different from primary storage)

[ ] Other: \_\_\_\_\_\_\_\_\_\_\_

### Data Security:

Are there any specific security considerations for your data, such as sensitive information?

[ ] No specific security concerns.

[ ] Yes (describe): \_\_\_\_\_\_\_\_\_\_\_

## 4. Data Sharing and Access

### Internal Sharing:

How will you share your data with your supervisor and other lab members?

[ ] Shared lab server with defined access permissions

[ ] Cloud storage with shared folders

[ ] Emailing files (for small datasets only)

[ ] Other: \_\_\_\_\_\_\_\_\_\_\_

### **External Sharing:**

Do you anticipate sharing your data outside the lab?

[ ] No external sharing anticipated at this stage.

[ ] Yes (please describe potential platforms, repositories or formats): \_\_\_\_\_\_\_\_\_\_\_

### Access Control:

Who will have access to your data, and how will you manage these permissions, especially on shared platforms?

[ ] Myself and my supervisor.

[ ] Specific lab members (list names/roles): \_\_\_\_\_\_\_\_\_\_\_

[ ] Access managed by lab IT or server administrator

## 5. Data Preservation and Re-use

### Data Retention:

How long will you keep your data after the project is complete? Are there any lab, journal or institutional guidelines you need to follow?

[ ] Following lab guidelines (please specify: \_\_\_\_\_\_\_\_\_\_)

[ ] Minimum of [Number] years.

[ ] Indefinitely.

### Data Archiving:

How will you archive your data for long-term preservation? Will you move it to a different storage location or format?

[ ] Moving data to a designated archive folder on the shared server.

[ ] Transferring data to an external hard drive for archiving.

[ ] Other: \_\_\_\_\_\_\_\_\_\_\_

### **Potential for Re-use:**

Do you see potential for your data to be re-used by yourself or others in the future? How will your data management practices facilitate this?

[ ] Yes, the well-documented and organized data could be useful for future projects or meta-analyses.

[ ] Potentially, if properly documented.

[ ] Unlikely.

## 6. Responsibilities

### **Data Management Responsibility:**

Who is primarily responsible for implementing and maintaining this data management plan?

[ ] The graduate student (you).

[ ] Shared responsibility with the supervisor.

### Training and Support:

What training or support do you need to effectively manage your data according to this plan? Who can you ask for help?

[ ] Guidance from my supervisor on best practices.

[ ] Assistance from lab members with specific techniques (e.g., file server usage).

[ ] Information on institutional data management policies.

Institutional and Other Resources

* [Open Science @ CAS](https://openscience.lib.cas.cz/en/)
* [FairWizard](https://fair-wizard.com/) (access through CAS)
* [Czech National Infrastructure for Biological Data](https://www.elixir-czech.cz/) (ELIXIR-CZ)
* [REPO national repository](https://data.narodni-repozitar.cz/) for Research Data <https://du.cesnet.cz/cs/navody/narodni_repozitar/start>
* Zenodo <https://about.zenodo.org/> - independent of field and institution
* <https://fairsharing.org/>
* <https://www.re3data.org/>
* <https://journals.plos.org/plosbiology/s/recommended-repositories>

## 7. Plan Review and Updates

### Review Frequency:

How often will you review and update this data management plan?

[ ] Every [Number] months.

[ ] At the start of each new major experiment.

[ ] As needed.

### Update Procedures:

How will changes to the plan be documented and communicated?

[ ] Updating this document and discussing changes with my supervisor.

[ ] Saving new versions of this document with a date.

### Signatures:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (Graduate Student)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (Supervisor(s))

Feel free to adapt this template to the specific needs and practices of your lab. The key is to be practical and establish habits that make your data findable, understandable, and reusable!

*This template was developed by Carol Wilusz (Colorado State University) using Gemini 2.0 Flash using prompt: “Suggest a data management plan template for a graduate student working at the bench in a molecular biology lab. This person does experiments in real time PCR, western blotting, microscopy, RNA sequencing and flow cytometry. They do not know how to code.” The template has been edited to provide resources for students at Czech Academy of Sciences. Feel free to personalize the prompt and generate your own template.*