

# General Guiding Principles for Data Sharing

## Data Format

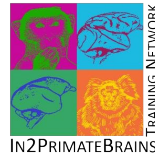
- The data curation process will involve archiving of original data and conversion to standardized format. Hence, all the shared data formats should be accompanied with:
  - Software/code used to read the data, and/or
  - Detailed documentation of the data format (links to the format description or the recording device might be helpful)

## Metadata

- It is crucial that the shared data is accompanied with as much metadata as possible.
- The metadata templates designed for the common localizer experiments are available in a dedicated repository on [GitHub](#) and on [GIN](#).
- In case of spreadsheets being used to provide additional information, use of the “Spreadsheet best practices” is encouraged to ensure that the information is machine readable. Please refer this page for the “Spreadsheet best practices”<sup>1</sup> - [https://ucdavisdatalab.github.io/workshop\\_keeping\\_data\\_tidy/#spreadsheet-best-practices](https://ucdavisdatalab.github.io/workshop_keeping_data_tidy/#spreadsheet-best-practices)

## File Organization

- The organization of the uploaded files must be consistent and well-documented.
- If the filenames and organization represents any metadata or semantic information, it must be documented. Few general guidelines for file names and folder names:
  - **Consistency:** Use a consistent naming convention across all files and folders.
  - **Clarity:** Names should be clear and descriptive, reflecting the content or purpose.
  - **No Spaces:** Avoid spaces in names; use underscores (\_) or dashes (-) instead.
  - **Include Identifiers:** Incorporate subject ID, date of collection, and other relevant identifiers.
  - **Versioning:** If multiple versions of a file exist, include a version number.
  - **Avoid Special Characters:** Do not use special characters like !, @, #, \$, etc.
  - **Extension:** Always include the appropriate file extension (e.g., .csv, .sev).
  - **Length:** Keep the names reasonably short while maintaining descriptiveness.
- Different task names included in the common localizer protocols could be standardized as follows:
  - task-reward6sManipulation
  - task-reward10sManipulation
  - task-videoMonkeyWorld123



- task-videoMonkeyWorld456
- task-videoMonkeyWorld789
- task-videoJsallet1
- task-videoJsallet2

Where the labels “reward” and “video” are two major categories of the tasks and the next label, such as “10sManipulation” represent the sub-category of the task.

- Below is a suggested directory structure that could be used for organizing your data:

```
|__project/
|  |__subjects.tsv
|  |__subjects.json
|  |__tasks.tsv
|  |__task.json
|  |__dataset_description.json
|  |__sourcedata/
|  |  |__sub-01/
|  |  |  |__sub-01_sessions.tsv
|  |  |  |__sub-01_sessions.json
|  |  |  |__ses-01/
|  |  |  |  |__ephys/
|  |  |  |  |  |__sub-01_ses-01_task-<name>.nix
|  |  |  |  |  |__sub-01_ses-01_ephys.json
|  |  |  |  |  |__sub-01_ses-01_channels.tsv
|  |  |  |  |  |__sub-01_ses-01_contacts.tsv
|  |  |  |  |  |__sub-01_ses-01_probes.tsv
|  |  |  |  |__beh/
|  |  |  |  |__sub-01_ses-01_eyetrack.tsv
|  |  |  |  |__sub-01_ses-01_camera.mp4
```

## Documentation

Include a README file ([template](#)) or equivalent in your submission, summarizing the key information about the data and metadata. This could include an overview of the data collected, a brief summary of methods used, and any other relevant details that will aid in understanding, curating and reusing the data.

If you have any questions or require some guidance, contact Reema (ESR14) at [reema.gupta@lmu.de](mailto:reema.gupta@lmu.de)