**PROJECT ORGANIZATION & DOCUMENTATION WORKSHEET**

#### 

#### **Instructions**

This worksheet is designed to assist you in organizing and documenting your capstone project while keeping reproducible and FAIR principles in mind. Please refer to the slides for examples and links we covered in class to review essential recommendations.

This exercise will allow your group to review and improve your current strategies. If your group started on the right track and you feel your systems are good enough, use this spreadsheet to describe the rationale behind your choices.

Each group should complete one worksheet (PARTS A-D) and a README.md file (PART E) describing your capstone project compendium. You may skip sections that do not apply to your project. Submit both files via Canvas.

#### **PART A:** Defining/revisiting your project needs

1. How do you currently organize your files? To answer the first question, you may take a screenshot of your capstone project directory (expanded) or use the 'tree' function in the command line, copy the code to Carbon (<https://carbon.now.sh/>), and paste the image here.
2. If there are any pre-established conventions from your client or how they would like your group to store and organize the files, please note them here.
3. Reflect on your current system and your recent use of it. Consider what works well and what doesn't work, and what you like and don't like about the system.

#### **PART B:** Creating/redefining your file naming schema

1. File inventory

Generate an inventory of all current files and list files you anticipate creating by the end of your capstone project.

Use the table below to capture the information on all your different file types.

* Column A: What different types of files are you reusing or creating? (e.g., vector, tabular, plots).
* Column B: What are the file formats? (e.g., .csv, .jpeg)
* Column C: What are the unique characteristics of this data file (e.g., date created, project name, experimental conditions)? Are there standard abbreviations for any of these characteristics? You may have multiple unique characteristics, including keywords, samples, locations, etc.
* Column D: Will you have multiple versions of the same file? If so, will there be various versions on a given day? How many?

| 1. Data Type | 1. File Format | 1. Unique characteristics | 1. Multiple Versions |
| --- | --- | --- | --- |
| E.g., Tabular | E.g., .csv | E.g.,  Water quality  Reservoir  Cachuma Lake  Morphology Data | E.g., One per year since 2010 |

If applicable, create a separate table to document any abbreviation, acronyms, or code used in the file naming.

| Descriptor | Abbreviation Keys | | |
| --- | --- | --- | --- |
| Location (County) | SB - Santa Barbara  CL - Contra Loma  Based on: <https://dot.ca.gov/-/media/dot-media/programs/local-assistance/documents/erp/2018/counties-abbr.pdf> | | |
| Analysis Type | MOR - Morphology  HYD - Hydrology | | |

2. Revise/Draft your file names

Copy the table above and add a fifth column for the file name.

Things to include: The unique characteristics you’ve identified in Column C can be combined to create a file name. Pick 4-5 elements you think will be helpful to identify and find these files easily.

If you have multiple versions of a file (indicated in Column D), add version

information to your file name. If you anticipate having more than one version in a day, you

can’t rely on the date alone and should add a version number to the file name. If that applies, stick to the [ISO 86010](https://www.ionos.com/digitalguide/websites/web-development/iso-8601/#:~:text=According%20to%20the%20basic%20format,%3A%2012%3A07%3A22.) standard.

Example: 20101231-MOR-Cachuma-SB.csv

#### **PART C:** Creating/redefining a systematic file folder hierarchy

1. Review your data types in the table in Part B. Think about how you would categorize these files. By file type? By time? By site? By project? By instrument?
2. Using these categories, determine a standard and consistent folder naming convention. These may mimic part(s) of your file names.
3. Consider the current hierarchy of your folders. How do folders relate to each other? Is this system working well for you (see A3)? If not, modify your system accordingly and sketch your revised hierarchy below.

#### **PART D:** Project Directory Organization

Based on your file inventory and the schema you have created to name your files. Now, describe your overall project directory organization (if different from Part A). Then, represent the hierarchical structure of your project directory, considering the recommendations provided and asking yourselves: Is your data file structure intuitive enough that a new project member or a future reuser could easily understand without asking you questions?

#### **PART E: README.md**

Based on your directory's revised/improved configuration, complete the [README.md template](https://drive.google.com/file/d/1LOOMvaUhMttFtQQvvxBvZ6QqlwwbZLiZ/view?usp=share_link) for your capstone project. Make sure to describe each file in your project directory, their relationships, and the variables and units of measurement they include (we will need these for Week 6 - Metadata).