

---

layout: page  
title: File Naming and Organization  
menu: true  
order: 1  
lang: English  
ref: filenaming

## permalink: /filenaming

---

Well-organized structures and consistent file names make it easier to keep track of and organize files, and are necessary for IDEP workflows. When naming and organizing files, you should be consistent and ordered so that files are easily identified and sort properly. Set up a clear directory structure and file naming convention that follows a pattern like `project identifier + grouping (like date or volume) + some type of unique identifier`.

## Structuring Data

Think carefully about how best to structure your data from the very earliest stages of your project. Try to keep a shallow hierarchy (no deep nesting of folders). Individual directories may be grouped by a number of different criteria, including date, format, sub-collection, or whatever categories make the most sense for your project.

Some examples of data structures include organization by:

- format (text, images, sound files, etc.)
- genre (interviews, posters, newspapers, postcards, etc.)
- unit (journal year/volume, archival box)

## File Names

File names should allow you to identify files with precision from the name. Choose a format for naming your files and use it consistently. Include in the name the information that will allow you to distinguish your files from one another, following a pattern like `project identifier + grouping (like date or volume) + some type of item identifier`. Notice how the elements move from general to specific. Some of the elements to consider include:

- Project or collection identifier: name, acronym, collection number, etc.
- Grouping: year, volume (journal), archival series or box, etc.
- Item identifier: page, issue (journal), sequence, unique identifier (ISBN, OCLC, shelf mark)

Some best practices:

- All special characters such as `~ ! @ - # $ % ^ & * ( ) ; < > ? , [ ] { } ' " |` should be avoided. Stick to numbers, Roman letters (preferably lower case), and underscores `_`.
- Do not use spaces or dashes. Some software will not recognize file names with spaces, and file names with spaces must be enclosed in quotes when using the command line.
- Alternatives to spaces include:
  - Underscores, e.g. `file_name.xxx`
  - No separation, e.g. `filename.xxx`
  - Camel case, where the first letter of each section of text is capitalized, e.g. `fileName.xxx`
- A good format for date designations is `YYYY_MM_DD` or `YYYY_MM`. This format makes sure all of your files stay in chronological order, even over the span of many years.
- Keep filenames as short as possible since long file names do not work well with all types of software. Include only the information that is necessary to find and make sense of the file. Aim to be meaningful and brief. Remember: the filename is not the place to record descriptive metadata!
- When using a sequential numbering system, use an adequate number leading zeros for clarity and to make sure files sort in sequential order. For example, use `0001`, `0002`, `...0010`, `0011` ... `0100`, `0101`, etc. instead of `1`, `2`, `...10`, `11` ... `100`, `101`, etc.

## Other Tips

Include a `README.txt` file in the directory that explains your naming format along with any abbreviations or codes you have used. This documentation will be helpful both during the project or experiment, and also in the future.