



File Naming Conventions

Karen Keehr, Archivist

October 2023

Considerations:

A file name is the chief identifier for a record. In the world of electronic records, the record's file name provides metadata that places the record in context with other records, records series, and record retention schedules. Smart file naming and considered choice of file formats will make a significant difference to the long-term preservation of digital records. Consistently using a standard file name format that also follows some basic rules will help ensure files can be found, opened, and preserved for as long as necessary. Choosing file formats that are robust and well-supported will aid the management and maintenance of growing numbers of digital records.

Records Will Be Accessed by Others:

Much of our work is stored on network servers and SharePoint so that they are accessible from multiple locations by various people. This requires that files names (as well as folder structures) make sense to more than just the files' creator. A file name should be clear to everyone at AHSGR. A record should be distinguishable from files with similar subjects as well as different versions of the same files.

Good Practices Quick Rules:

- Never use spaces or special characters like: !*&%
- Do use dashes (-) and underscores (_) between words or phrases
- Never change the file extension through re-naming
- Use a short descriptive file name of content and date
- Include all necessary descriptive information independent of where it is stored
- Use a standard date format: YYYY_MM_DD
 - This date format is best for sorting digital records
- Use a consistent method for showing the file versions, such as v1
- For any files from repeating events, put date at the end
 - For example, staff meeting agendas: Staff_Meeting_Agenda_2023_06_07
- BE CONSISTENT, but there will be exceptions

File Naming Rules Breakdown:

Rule #1: Avoid using special characters in a file name such as \ / : * ? " < > | [] & \$, .

Computers and other electronic devices frequently use the characters listed above for specific tasks in an electronic environment. For example, a forward slash is used to identify folder levels in Microsoft products, while Mac operating systems use the colon. Periods are used in front of file-name extensions to denote file formats such as .jpg and .doc; using them in a file name could result

in lost files or errors. Using these reserved characters can cause the system to misinterpret or be unable to read the file.

Rule #2: Use underscores instead of periods or spaces

As mentioned above, periods already have a specific function in a file name, which is to tell the computer program where the file-name extension begins. Spaces are frequently translated in a web environment to be read as “%20”. For example, if it were available online,

Naming tutorial.doc

would appear as

Naming%20tutorial.doc

This alteration can cause confusion in identifying the actual file name. Spaces in file names can also cause broken links, because word processing tools like Microsoft Word, and email clients like Microsoft Outlook, recognize spaces as an opportunity to move to another line. Therefore, a link to

\\Ah1\Intranet\ar\naming conventions

could become

\\Ah1\Intranet\ar\naming
Conventions

Rule #3: Err on the side of brevity

Different operating systems and software have different limits on the acceptable length of file names and file paths. Some systems allow up to 256 characters, while others allow far fewer.

Keeping file names short can help prevent any future issues. Generally, about 25-30 characters is a sufficient length to capture enough descriptive information for naming a record.

Rule #4: The file name should include all necessary descriptive information independent of where it is stored

Files are frequently copied to other folders, downloaded, and emailed. They may be migrated to newer storage or moved permanently if transferred to the archive. It is important to ensure that the file name, independent of the folder where the original file lives, is sufficiently descriptive.

Electronic records are usually organized in a series of folders. For example:

World_War_I\Posters\Owens\0001.tif

While this is a very organized way of storing records, it is only efficient as long as the files stay in their original folders in their original context. When multiple staff work on a project or if staff are multitasking, it is very easy to misfile a document. As soon as “0001.tif” is copied to another folder, or emailed to an agency, the context provided by the folders in which the document is nested is lost. Context is particularly important in legal situations because it provides authenticity and trustworthiness of the record. Losing the context of a record can possibly compromise its trustworthiness and therefore its validity or admissibility for court proceedings. Additionally, if the file has the same name as another file on the system in a different folder, once it is pulled out, the computer could mistake it for the same file and write over one of the files.

For instance, if the following files were pulled out of their appropriate folders, they would appear to be the same file:

World_War_I\Posters\Owens\0001.tif
World_War_I\Posters\RedCross\0001.tif
Incorrect: 0001.tif
Correct: wwI_poster_owens_0001.tif

Rule #5: Include dates and format them consistently

The point of organizing the agency's electronic records is to enable accessibility not only by current users, but by future users as well. Records retention schedules are applied to electronic records just as they are to paper records. Having the relevant date associated with the file is essential. Though many operating systems store this information with the file, as users move the file among folders and computers and as the file is re-saved as revisions are made, those dates change. A file could have a "created on" date that does not accurately reflect when they were created.

The best way to prevent confusion is to embed the relevant date (the date that the file was created or revised) in the file name itself.

For consistency, put the date at the end of the file name. For example:

File_Name_YYYY_MM_DD

The best way to list the date is based on an international standard – ISO 8601. ISO 8601 specifies numeric representations of date and time to be used in electronic format. The international standard date notation is:

YYYY_MM_DD

YYYY is the year, MM is the month of the year between 01 (January) and 12 (December), and DD is the day of the month between 01 and 31. For example, January 5, 2008 is written as 2008_01_05.

This format allows ease of sorting and comparing files by date and prevents confusion with other date formats (especially in other formats that use just two digits for the year).

For example, this document could be named:

filenaming_2008_05_07

to reflect that this draft was last edited on May 5, 2008.

Rule #6: Include a version number on documents to manage drafts and revisions more easily

A file will frequently have multiple versions, especially when it is created by a workgroup. Specifying the version of a file can help you quickly identify the most accurate or most definitive version of the document.

The easiest way to do this is to use the letter "v" to represent "version number." Then, "v01, v02, v03" can be added as needed to a file and the main file name can stay the same. This is much more effective than other common additions like "update," "new," "old," etc. An exception to this rule is

using “FINAL” to indicate the final version of the document. When using “FINAL”, be sure to use it *instead* of the version number, rather than in addition to it.

Rule #7: Be consistent.

The most important rule of file naming is to be consistent. Some choices will need to be made about organizations that affect the entire workgroup – where to include the date, what abbreviations to use, etc. Regardless of what the group decides, it is only effective if everyone follows the rules consistently.

There Will Be Exceptions

One notable consideration is the batch-scanning process. This process typically relies on a program that sets its own parameters on the file names allowed. If this is the case, take advantage of the folder hierarchy and, when possible, apply the rules outlined here to folder names. Avoid spaces and special characters. Be consistent throughout the project; consider developing a file-naming standard for all batch-scanning projects.

Remember: *this document is not going to apply absolutely to every situation; it should be used as a guide to encourage discussion in offices and workgroups to develop file-naming practices that work best in those specific environments.*