Digital Media File Naming Guide

About this guide

The media files that are created in your projects must be given appropriate and consistent file names and held in appropriate storage locations. See the appropriate section below for information about naming conventions, and folder structures for media files created by digitizing different media formats.

General file naming practices

There are a few elements that you will want to include to create effective file names for your digital assets, and there are a few things to avoid. The following are helpful practices in naming your digital files:

- keep file names relatively short
- separate words in file names
- use leading zeroes to keep your files sorting properly
- avoid the use of most non-alphanumeric characters

For more information on these file naming practices, please refer to the ==*File Management Guide*==.

Naming conventions for digitized media files

Typically, the following digital media files are created during digitization:

- preservation file (pf)—a high quality digital copy of an analog audio, or video tape. This file will have little or no alterations in order to act as a faithful representation of the original tape. These files are generally quite large compared to other types of files.
- *primary access file* (paf)—a high quality copy of the *preservation file* intended for access by a person who is allowed to listen to, or view the content. This file may be edited to improve the visual or audible clarity of the media. You may also edit, or redact, sensitive information to allow for different levels of access.
- *access copy* (ac)—you may wish to create copies of the *primary access files* at a variety of qualities to allow for differing methods of distribution (i.e. email, flash drive or Internet download).
- subclips—subclips media files that contain portions of the content of files listed above. These files are generally created when that content is requested and can be ephemeral or managed with your digital media collection.

When naming digitized media files, you should follow a predictable and consistent pattern. It is a requirement that any audio or video tapes will have been given a **unique identifier** before

digitization. The file names of these digitized tapes should be based on their unique identifier. This connects the media file to the physical object.

For more information on unique identifiers, please refer to the <mark>Describing and Managing Media Collections Guide</mark>. The examples of unique identifiers used below are from that guide.

Other information included in the file name will vary depending on the media. It is not quite enough to connect the digitized file to the physical tape. If there are several recordings on a type of media then the filename should also describe what part of that tape is represented in the digital file. For example, audio cassettes often have recordings on both" side a" and "side b" or the tape. It is not enough to know that a digitized audio file is from a specific tape. You also need to know if it from side a or b of that tape. File names for open reel audio recordings can become much more complicated to create as there is a lot of variability in the way that open reel decks record audio signals onto the tape.

Information about recordings is sometimes called **metadata**. As mentioned above, you will need to capture different **metadata** in the file name depending on the media type. Examples of information that might be tracked in the filename are what **side**, physical **track**(s)/channel(s) or recorded **signal** are represented in the files. Also, you can embed the intended use of that media file into the file name (i.e. **preservation file "pf"**, **primary access file "paf"** or an **access copy "ac"**).

- Separate unique identifiers and other metadata elements with underscores to make file names easier to read.
- If sub-copies or clips/portions of access copies are created, a sequential number may be added to the end of the file name (i.e. cc0001-1a_ac_001.mp3)
- The file name is followed by a file extension (such as .wav, .mp3, .avi or .mp4). These file extensions are typically populated by software programs and do not need to be typed into a file name.

Note: Do not worry if you have already been digitizing media and the suggested file naming schemes do not match what you have been using. You can continue with existing naming conventions or adjust for future digitization.

File names for digitized audio cassettes

Audio cassettes are very standardized in the way that they record audio information onto a tape. Typical audio cassettes record on four physical tracks (strips) on the magnetic tape. It is very consistent which two tracks that are read when playing a given "side" of the cassette. For this reason, we include what side of the cassette is represented in the file name but not what physical track was played (unlike open reel audio recordings). Subclip information can be included in file naming for any format, and generally is only used for access files.

Example file naming convention:

Audio cassettes were created by the Phillips Corporation who named the format "Compact Cassette". Even though the format is most commonly known as "audio cassettes", the Indigitization Program often uses the acronym "cc" in relevant unique identifiers and file names. This is because we already use the abbreviation "ac" in file names to designate "(a)ccess (c)opy".

See Audio Cassette File Naming Chart in Appendix for more examples

File names for open reel audio

The open reel format is also referred to as "reel to reel". Although we use the term open reel in our guides, we generally use the acronym "rr" in unique identifiers and filenames. This is due to the fact that the letter "o" and the number zero ("0") can sometimes look very similar depending upon the font that it is represented in. It is for the same reason that people often avoid the lower case L ("l") in identifiers, as it can sometimes look similar to a number one ("1").

Open reel audio tapes have a wide range of track patterns that can be recorded on the magnetic tape. These tapes can also be played in one of two physical orientations (side a and side b). The open reel format also supports many recording speeds, which must be set manually (as opposed to video formats where speed is automatically changed). Differences in recording speed can require open reel tapes to be digitized in multiple sessions to properly accommodate these differently recorded "signals". For this reason, we recommend including information in the file name about what side of the reel is represented, what tracks were played, and which signal was captured (if different parts of the tape had to be digitized in different sessions). Subclip information can be included in file naming for any format and generally is only used for access files.

Where recording collections are very consistent and follow the most common standards, you can decide to simplify open reel filenames to be similar to that of audio cassettes. An example of such a case would be a collection of tapes that are all recorded in a standards side a/b track pattern, and at a single speed for each entire "side" of the tape.

See Open Reel File Naming Chart in Appendix for more examples

Appendix

Audio cassettes

Example unique identifier: **cc0001-1** (**c**ompact **c**assette #**0001**, part**1**)

FILNAME ELEMENTS						
Unique identifier (Tape ID)	**Side**	**Preservation File (pf) Primary Access File (paf) Access Copy (ac)**	**Subclip**	**Extension**	**Filename(s)**	**File description**
cc0001-1	a	pf paf ac	001 002 003	.wav .mp3	**cc0001-1a_pf.wav**	side a - preservation file
					cc0001-1b_pf.wav	side b - preservation file
					cc0001-1a_paf.wav	side a - primary access file
					cc0001-1b_paf.wav	side b - primary access file
					cc0001-1a_ac.mp3	side a - access copy
	b		none		**cc0001-1b_ac.mp3**	side b - access copy
					cc0001-1a_ac_001.mp3	side a - access copy – subclip #1
					cc0001-1a_ac_002.mp3	side a - access copy – subclip #2
					cc0001-1a_ac_003.mp3	side a - access copy – subclip #3

The table above represents an example set of filenames for a digitized cassette using a simple tape identifier. In this case we are creating a typical set of files (Preservation File, primary access copy and an access copy) for each side. There are also three subclips for side a that are only created as access copies.

Example unique identifier (with sub-collection identifier): **A2021-01-cc01**
(sub-collection identifier **A2021-01**, **c**ompact **c**assette **#01**, part**1**)

FILNAME ELEMENTS						
Unique identifier (Tape ID)	**Side**	**Preservation File (pf) Primary Access File (paf) Access Copy (ac)**	**Subclip**	**Extension**	**Filename(s)**	**File description**
A2021-01-cc01-1	a	pf paf ac	001 002 003	.wav .mp3	**A2021-01-cc01-1a_pf.wav**	side a preservation file
					A2021-01-cc01-1b _pf.wav	side b preservation file
					A2021-01-cc01-1a _paf.wav	side a primary access file
					A2021-01-cc01-1b _paf.wav	side b primary access file
					A2021-01-cc01-1a _ac.mp3	side a access copy
	b		none		**A2021-01-cc01-1b_ac.mp3**	side b access copy
					A2021-01-cc01-1a _ac_001.mp3	side a access copy – subclip #1
					A2021-01-cc01-1a _ac_002.mp3	side a access copy – subclip #2
					A2021-01-cc01-1a _ac_003.mp3	side a access copy – subclip #3

The table above represents an example set of filenames for a digitized cassette using a tape identifier that incorporates a sub-collectionID. In this case we are creating a typical set of files (Preservation File, primary access copy and an access copy) for each side. There are also three subclips for side a that are only created as access copies.

Open reel audio

Example unique identifier: **rr0012-1** (reel to reel #0012, part1)

Note: We are using "rr" instead of "or" as "o" can be mistaken for "0" (zero) in some fonts. Many managers choose to use easily readable characters in their file names.

FILE ELEMENTS							
Unique identifier	Side	Tracks	Signal	Preservation file (pf) primary access file (paf) access copy (ac)	Subclip	Filename(s)	File description
rr0012-1	a	t13	s1	pf paf ac	001 002	**rr0012-1a_t13_s1_pf.wav**	side a preservation file (track 1 & 3) signal #1 (only signal)
						rr0012-1b_t42_s1_pf.wav	side b preservation file (track 4 & 2) recorded signal #1 (signal 1 of 3)
						rr0012-1b_t42_s2_pf.wav	side b preservation file (track 4 & 2) recorded signal #2 (signal 2 of 3)
						rr0012-1b_t42_s3_pf.wav	side b preservation file (track 4 & 2) recorded signal #3 (signal 3 of 3)
						rr0012-1a_t13_s1_paf.wav	side a primary access file
						rr0012-1b_t42_s1_paf.wav	Side b, signal #1 primary access file
						rr0012-1b_t42_s2_paf.wav	Side b, signal #2 primary access file
	b	t42	s1 s2 s3		none	**rr0012-1b_t42_s3_paf.wav**	Side b, signal #3 primary access file
						rr0012-1a_t13_s1_ac.mp3	side a access copy
						rr0012-1b_t42_s1_pf.mp3	side b access copy, signal #1
						rr0012-1b_t42_s2_pf.mp3	side b access copy, signal #2

						rr0012-1b_t42_s3_pf.mp3	side b access copy, signal #3
						rr0012-1a_t13_s1_ac_001.mp3	side a access copy – subclip #1
						rr0012-1a_t13_s1_ac_002.mp3	side a access copy – subclip #2

Example unique identifier (with sub-collection identifier): **A2021-04-rr01-1**
(sub-collection identifier A2021-04, reel to reel #01, part1)
Note: We are using "rr" instead of "or" as "o" can be mistaken for "0" (zero) in some fonts. Many managers choose to use easily readable characters in their file names.

FILE ELEMENTS						Filename(s)	File description
Unique identifier	Side	Tracks	Signal	Preservation file (pf) primary access file (paf) access copy (ac)	Subclip		
A2021-04-rr01-1	a	t13	s1	pf paf ac	001 002	A2021-04-rr01-1a_t13_s1_pf.wav	side a preservation file (track 1 & 3) signal #1 (only signal)
						A2021-04-rr01-1b_t42_s1_pf.wav	side b preservation file (track 4 & 2) recorded signal #1 (signal 1 of 3)
						A2021-04-rr01-1b_t42_s2_pf.wav	side b preservation file (track 4 & 2) recorded signal #2 (signal 2 of 3)
						A2021-04-rr01-1b_t42_s3_pf.wav	side b preservation file (track 4 & 2) recorded signal #3 (signal 3 of 3)
						A2021-04-rr01-1a_t13_s1_paf.wav	side a primary access file
						A2021-04-rr01-1b_t42_s1_paf.wav	Side b, signal #1 primary access file
						A2021-04-rr01-1b_t42_s2_paf.wav	Side b, signal #2 primary access file
	b	t42	s1 s2 s3		none	A2021-04-rr01-1b_t42_s3_paf.wav	Side b, signal #3 primary access file
						A2021-04-rr01-1a_t13_s1_ac.mp3	side a access copy
						A2021-04-rr01-1b_t42_s1_pf.mp3	side b access copy, signal #1
						A2021-04-rr01-1b_t42_s2_pf.mp3	side b access copy, signal #2
						A2021-04-rr01-1b_t42_s3_pf.mp3	side b access copy, signal #3
						A2021-04-rr01-1a_t13_s1_ac_001.mp3	side a access copy – subclip #1
						A2021-04-rr01-1a_t13_s1_ac_002.mp3	side a access copy – subclip #2

For details on sides, signals, and tracks, please refer to the ==Open Reel Digitization Guide==

Video formats

Example unique identifier: **VHS0001-1** (VHS tape #0001, part1)

Unique identifier (Tape ID)	Preservation file (pf) primary access file (paf) access copy (ac)	Subclip	Extension	Filename(s)	File description
VHS0001-1	pf paf ac	001 002 003	.avi .mp4	VHS0001-1_pf.avi	preservation file
				VHS0001-1_paf.avi	primary access file
				VHS0001-1_ac.mp4	access copy
				VHS0001-1_ac_001.mp4	access copy – subclip #1
				VHS0001-1_ac_002.mp4	access copy – subclip #2
				VHS0001-1_ac_003.mp4	access copy – subclip #3

Example unique identifier (with sub-collection identifier): **A2021-01-vhs01-1**
(sub-collection identifier A2021-01, VHS tape #01, part1)

Unique identifier (Tape ID)	Preservation file (pf) primary access file (paf) access copy (ac)	Subclip	Extension	Filename(s)	File description
A2021-01-cc01-1	pf paf ac	001 002 003	.avi .mp4	A2021-01-VHS01-1_pf.avi	preservation file
				A2021-01-VHS01-1 _paf.avi	primary access file
				A2021-01-VHS01-1 _ac.mp4	access copy
				A2021-01-VHS01-1 _ac_001.mp4	access copy – subclip #1
				A2021-01-VHS01-1 _ac_002.mp4	access copy – subclip #2
				A2021-01-VHS01-1 _ac_003.mp4	access copy – subclip #3

Further examples:

Unique identifier (Tape ID)	Preservation file (pf) primary access file (paf) access copy (ac)	Subclip	Extension	Filename(s)	File description
8mm001-1	paf		.avi .	8mm001-1_paf.avi	primary access file

Unique identifier (Tape ID)	Preservation file (pf) primary access file (paf) access copy (ac)	Subclip	Extension	Filename(s)	File description
c002-vhsc01-1	ac	002	.mp4	c002-vhsc01-1_ac_002.mp4	access copy – subclip #2