Purpose

The purpose of establishing a file naming standard for material digitized by Yukon Archives is to:
  - ➢ allow for the application of unique identifiers to digital files;
  - ➢ allow for readability of file names (human readable)[1];
  - ➢ allow for descriptive file names;
  - ➢ allow for the association of digital files with their source material;
  - ➢ minimize duplication of file names for different source material;
  - ➢ minimize decisions required to name a file;
  - ➢ allow for ease in searching for a file;
  - ➢ allow for an organized structure of storage;
  - ➢ maximize portability; and
  - ➢ minimize problems with exchanging files.

The conventions defined here have been developed and agreed upon by Yukon Archives staff to meet the above objectives. They are specific to Yukon Archives.

Standard Naming Convention

The following are specific rules about the formation of individual components in the file name:

  - ➢ A file name is composed of a base name and an extension, separated by a period.

  - ➢ File name elements correspond to those used with the physical materials.

  - ➢ Base names are composed of ASCII numerals (0-9), ASCII letters (A-Z, a-z) and the underscore (_). No other characters are used including spaces.[2] Case sensitivity is not assumed (i.e. PAM, Pam, and pam are considered to be the same).

---

[1] The convenience of meaningful file names makes it simpler and quicker to manually manage folders and files.
[2] Although some file naming standards do not allow for lowercase letters and other standards allow for a wider range of characters, this character set has been identified to best achieve the purposes outlined.

- ➤ Extensions are determined by the file format of the digital file. They are not arbitrarily assigned (i.e. JPG or TIF and not ABC or XYZ).

- ➤ Recommended maximum file name length is 30 characters.[3]

- ➤ Accession numbers[4] are uppercase. Location codes are uppercase, while other location references are lowercase. Descriptions follow standard capitalization rules.

- ➤ Use 4 digits for accession years 2000 onward (i.e. 2001 not 01).

- ➤ The underscore (_) separates some base name elements and words in descriptions and replaces the slash (/) in accession numbers. Characters such as apostrophes, colons, and hyphens in original titles can simply be omitted.

- ➤ Do not use 0 placeholders except for catalogued photographs (e.g. 004519) and sequential PAM numbers (e.g. 0045). Note that the base number of a Dewey Decimal Classification number always has three digits (e.g. 016) and this is reflected in these naming conventions as well.

- ➤ Dates are written in the ISO 8601 basic format YYYYMMDD (e.g. 19020305 for Mar 5th 1902) to allow for natural sorting. If the day is not required, use YYYYMM (e.g. 190203 for March 1902).

- ➤ File name elements are more detailed / specific as they are read from left to right.

- ➤ When an entire grouping of material is digitized, each digital file is given a sequential number.

- ➤ Occasionally, additional elements (e.g. volume or edition) or deviations are required due to exceptions encountered with the physical material. These are addressed on a case by case basis, and generally occur in the more detailed / specific portion of the file name.

- ➤ File name elements should not contain information that might change over time.

---

[3] Although ideal, Yukon Archives will not require an 8.3 file name standard (8 characters for the base name and 3 characters for the extension) due to the difficulty in capturing the necessary information to allow for the purposes outlined.
[4] acc year = 2 or 4 digit accession year portion of accession number
acc no = incremental number portion of accession number
acc no code = letter code portion of accession number if present (i.e. R).

Photographs:

Catalogued:

FORMAT : YA# YA# suffix _ description if necessary

000136.tif	*six digits*
001481A.tif	*YA# suffix at end*
008946_back.tif	*description at end (do not name the front 008946_front.tif)*

Uncatalogued:

FORMAT : acc year _ acc no acc no code _YA# prefix _ YA # _ YA# suffix _ description if necessary

82_321_46.tif	*no spaces*
2002_118_3.tif	*use 4 digits for accession years 2000 onward*
2005_3_5.tif	*do not use 0 placeholder on accession number or on photo number*
91_51R_35.tif	*with acc no code,* R *uppercase*
82_563_cs706_11.tif	Photographs numbered by contact sheets:
	YA# prefix, cs *lowercase*
82_563_f10_1.tif	Photographs numbered within folders:
	YA# prefix, f *lowercase*
82_328_1B.tif	*YA# suffix at end (upper or lowercase depending on how catalogued)*
99_68_187_back.tif	*description at end (do not name the front 99_68_187_front.tif)*
82_346_417to82_346_418.tif	*all photos from __ to __ (do not use 'and' in place of 'to')*

Albums:

FORMAT: acc year _ acc no acc no code _ page # or description

FORMAT for ENTIRE ALBUMS: acc year _ acc no acc no code _ location code location code counter location code counter code _ file # _ object placeholder sequential number

2009_99_pg31.tif	*page 31*

2009_99_PHO_OS_27_f1_31.tif   *31st object in sequence for entire file*

do not use 0 placeholder on sequential number (i.e. not _031)

Textual Records:

FORMAT: acc year _ acc no acc no code _ location code location code counter location code counter code _ file # _ description (subject / keyword(s) / date, <u>from</u> and date for correspondence) or page # if file is small

FORMAT for ENTIRE FILES: acc year _ acc no acc no code _ location code location code counter location code counter code _ file # _ object placeholder sequential number

Manuscripts:

93_5_MSS45_pg1.tif	*do not use 0 placeholder on location code counter (i.e. not 045)*
2000_3_MSS407_*subject*.tif	*no spaces, no underscore after MSS*
86_49_MSS197_typed_will.tif	*MSS uppercase*
81_140_MSS171_f8_18.tif	*18th object in sequence for entire file*
	do not use 0 placeholder on sequential number (i.e. not _018)

Corporate Records:

79_63_COR86_f5_pg56.tif	*do not use 0 placeholder on location code counter (i.e. not 086)*
2001_30_COR999_f15_*subject*.tif	*no spaces, no underscore after COR*
88_128_COR998_f3 _Beck_19720412.tif	*use file#, <u>from</u> and date for correspondence*
88_128_COR307_f15_*keyword(s)_from_item_title*_19700103.tif	*use file#, keyword(s)/ description/subject and date*
82_454_COR224_f8_38.tif	*38th object in sequence for entire file*
	do not use 0 placeholder on sequential number (i.e. not _038)

Government Records:

GOV1003_f3_pg5.tif

if no accession number, if file is small

GOV100_f11_Gibbons _19340412.tif

use from and date for correspondence

GOV100_f11 _Gibbons _19340414_pg2.tif

use page# if more than one page in correspondence

GOV371_188.tif

if file contains numbered documents

97_90_GOV3237_f3 _Kimmerly_19861209_pg1.tif

if accession number

83_24_GOV1182_f1_4.tif

4th object in sequence for entire file

do not use 0 placeholder on sequential number (i.e. not _004)

Books:

FORMAT: title _ page # or description

FORMAT for ENTIRE BOOK: title _ object placeholder sequential number

Yukon_Manhunt_cover.tif   *title (or portion) and page*

Yukon_Manhunt_pg45.tif

Yukon_Manhunt_45.tif   *45th object in sequence for entire book*

do not use 0 placeholder on sequential number (i.e. not _045)

Pamphlets:

FORMAT: PAM # _ description

FORMAT for ENTIRE PAM: PAM # _ object placeholder sequential number

PAM1898_0005_pg3.tif

no spaces, no underscore after PAM, PAM all uppercase

PAM1954_0013_pg8to9.tif

example of 2 pages scanned as one file

PAM1924_0002_SSCanadian.tif   *description at end*

PAMND_0250C_pg9.tif

Coutts collection pamphlet with no date

PAM1898_0005_3.tif

3rd object in sequence for entire PAM

do not use 0 placeholder on sequential number (i.e. not _005)

Newspapers:

FORMAT: title _ date _ page # or description

FORMAT for ENTIRE NEWSPAPER: title _ date _ object placeholder sequential number

YukonNews_19650315_pg3.tif

WHSE_19550919_*subject*.tif

Whitehorse Star abbreviation

DDN_192301_87.tif

87th object in sequence for entire newspaper month

do not use 0 placeholder on sequential number (i.e. not _087)

Dawson Daily News Abbreviation

DWN_19430104_7.tif

7th object in sequence for entire newspaper day

Dawson Weekly News Abbreviation

ChooutlaGrayling_1955_1956_3.tif

community newsletter with date of 1955-1956

Periodicals:

FORMAT: media _ title _ date _ page # or description

FORMAT for ENTIRE PERIODICAL: media _ title _ date _ object placeholder sequential number

Periodical_CaribooNWDigest_1948_Summer_pg127.tif

not issued monthly

Periodical_CaribooNWDigest_1948_Summer_127.tif

127th object in sequence for entire periodical

Posters / Maps:

FORMAT: media _ location code _ location code counter _ description if necessary

Posters:

Poster_P_34.tif	*precede location code with* Poster_
Poster_P_230_bottom.tif	*description at end*

Maps:

Map_H_305.tif	*precede location code with* Map_
Map_H_2679_shF.tif	*use sheet number/letter when necessary*
Map_F_3.tif	
Map_R_36.tif	
Map_H_OS_2.tif	*oversize horizontal maps*
Map_ROLL_3_*description*.tif	
Atlas_A_77_pg49.tif	*precede location code with* Atlas_

Original Art / Miscellaneous:

FORMAT: media _ location code _ location code counter _ description if necessary

OrigArt_O_8.tif	*precede location code with* OrigArt_
Misc_M_24_top.tif	*precede location code with* Misc_
	description at end

Sound Recordings:

PRESERVATION FORMAT: location code location code counter _ series # _ side

SR33_1_A.wav	*letter for side*[5]
	SR 33 (1) first side

PRODUCTION FORMAT: location code location code counter _ series # _ side item

SR33_1_A1.wav	*letter for side, number for item*
	SR 33 (1) first side first item
SR35_2_B4.wav	*SR 35 (2) second side fourth item*

---

[5] If it is necessary to create multiple preservation master files for one side (i.e. recorded at different speeds or with wide differences in recording levels, damaged, etc.), use the file naming standard for production formats and consider each digital file as an 'item'. Note: In this case, although a preservation and production file may have the same file name, they could have completely different content.

SR84_1_B3toSR84_2_A2.wav	*all content from SR 84 (1) second side third item to SR 84 (2) first side second item (do not use 'and' in place of 'to')*
PHONO11_1_B2.wav	*PHONO 11 second side second item*
CD47_3_A4.wav	*CD 47 (3) first side fourth item*
	for CD/DVD Audio

Moving Images:

PRESERVATION FORMAT: location code location code counter

PRODUCTION FORMAT: location code location code counter _ item

Film:

to be determined

Video:

V5.mov	*preservation format[6]*
	V-5
V705_2.mov	*production format*
	V-705 second item

In Process Material:

FORMAT: as per examples above but with _temp and any other relevant information for that media added

94_58_temp_box8_f18_2.tif	*temporary uncatalogued photograph number*
	accession number followed by _temp followed by other relevant information (lowercase)
V1202_temp_34.mov	*temporary video number*
	location followed by _temp followed by other relevant information
	temporary number for the 34th item in a grouping

---

[6] If it is necessary to create multiple preservation master files for one video (i.e. recorded differently, damaged, etc.), use the file naming standard for production formats and consider each digital file as an 'item'. Note: In this case, although a preservation and production file may have the same file name, they could have completely different content.

Derivative Standard Naming Convention

Either of the following two methods can be used when naming derivative files if the derivative is of the same extension as the master:

> ➢ Preferred Method: use the same file name and create a series of folders that are named for the type of derivatives to indicate the intended use of the files (i.e. web files, full size, full screen, thumbnails); or
> ➢ Alternate Method: add descriptive suffixes to each base name to indicate the intended use of the files (i.e. _ac Access, _der for derivative, _fs for full-size, _sc for full screen, _tn for thumbnail, _sb for sound bite).

If a derivative file is created by combining multiple files, add or remove elements of the file naming convention as required (e.g. 81_140_MSS171_f8.pdf for all objects in an entire manuscript file, Yukon_Manhunt.pdf for all objects in an entire book, SR84_1_B3toSR84_2_A2.wav for two sound recordings combined).

If copies of Production and Preservation Master Files need to be copied and differentiated from one another add _pm (Preservation Master) or _pd (Production Master) to the end of each base name.

Associated Files Standard Naming Convention

For the base name of files associated with a specific digitized asset, use the base name and extension of the digitized asset separated by an underscore (e.g. 82_321_46_tif.txt, V5_mov.xml).

Folder Standard Naming Convention

The following are specific rules about the formation of folder names:

> ➢ Names are composed of ASCII numerals (0-9), ASCII letters (A-Z, a-z) and the underscore (_). No other characters are used including spaces. Case sensitivity is not assumed.

> ➢ Recommended maximum folder name length is 30 characters.

> ➢ When composing folder names, incorporate file name element formats (e.g. 82_454_COR224_f8, WHSE_194301). When composing folder names for books incorporate the Dewey Decimal Number (e.g. Ref_551_483_YRBC_Sept_1984_*title*, F_Wie_2003_Mad_Trapper)

> ➢ Recommended maximum folder hierarchy is eight levels.

> ➢ Recommended maximum path length is 255 characters.

<u>Original signed by David Schlosser</u>   <u>2018-01-22</u>
Territorial Archivist   Date

1st revision   2005-12-07
2nd revision   2006-03-27
3rd revision   2008-05-09
4th revision   2009-05-25
5th revision   2013-01-07