File naming conventions for Collections

Before you begin to scan and name files for your MMP collection(s), decide how you would like to name your files. This will ensure consistency, uniqueness in file names, and efficiency.

It is important to know the scope and details of the collection, prior to establishing a file naming convention. Take the size of the collection into account; the naming convention should allow for unique file names for each item.

Required:
1. All filenames should be unique
2. The file-naming schema should provide a logical sequence that can be easily identified and reproduced
3. Use only alpha-numeric characters for both files and folders.
   - Exceptions are dashes (-) and underscores (_)
4. Do not use special characters.
   - Examples include: /><+=^{}[]#
5. Use a three-character file extension (i.e. ".tif," ".pdf", etc.) at the end of each file name
6. Do not use spaces in file/folder names—use underscores instead
7. All letters should be lower case (Some operating systems are case sensitive. Using lower case consistently prevents problems.)
8. Use leading zeros
   - When file names need numbering, use zeros as placeholders. Example: a collection of 900 items should be numbered: abc001.tif, abc002.tif, etc.
     o This will facilitate proper sorting and file management
9. When dates are used, base the date on the MMP standards
   - Example: yyyymmdd, 20121030, or 2012_10_30
10. File names should contain necessary descriptive information, independent of their storage location.
    - Examples: For a Helena Independent Record newspaper from February 13, 1910:
      o hir1910_02_13_001.tif (Helena Independent Record; February 13, 1910 Page 1)
      o hir1910_02_13_002.tif (Helena Independent Record; February 13, 1910 Page 2)

Naming tips:

File pages will be alpha-numerically sorted
Daily Herald, April 1, 1900 (folder name)
      0001_Title page.tif
      0002_Masthead.tif
      0003_Page1.tif

Will appear in the navigation pane as:
Daily Herald April 1, 1900
      Title Page
      Masthead
      Page 1

MMP Newspaper File Naming Sample:

Flathead Courier 1910-01-01.001
Flathead Courier 1910-01-01.002
Flathead Courier 1910-01-01.003
Flathead Courier 1910-01-01.004

Flathead Courier 1910-01-08.001
Flathead Courier 1910-01-08.002

Flathead Courier 1910-01-08.003
Flathead Courier 1910-01-08.004

Flathead Courier 1910-01-15.001
Flathead Courier 1910-01-15.002
Etc.

It is possible that the file name may someday appear as a URL and that may be why the recommendations always provide a file name without spaces. So here's my latest suggestion on how to name these files - and it actually conforms to the current suggestion in the MMP Guidelines with the small change of adding underscores to make it more readable on the viewer end.

For the Flathead Couriers, how about this: FC1910_01_01_001.tif ...and so on. That way we have a file name that shows the date in a more readable way and also a file name with no spaces. Some librarians might want to create their own descriptive file names when they upload into the project and that would be alright but this might be a good convention for those who don't have time. What do you think?

Sources referred to:
http://ucblibraries.colorado.edu/systems/digitalinitiatives/docs/filenameguidelines.pdf
https://digitalriceprojects.pbworks.com/w/page/35749942/File%20naming%20conventions
http://www.library.illinois.edu/dcc/pdfs/best_practicespdfs/02_best_practices_for_file_naming_opt.pdf
http://library2.usask.ca/spcoll/Content%20DM%20tutorials/Working%20With%20Content%20Types/11managing-newspapers.pdf