5 Tips For What NOT To Do When Creating A File Naming Structure

1. **Do not develop a file name dependency**
   File names are not actually part of the file, but rather part of the file system and are therefore not dependable as being persistent over time and across systems. The Unique ID (UID) assigned to the object should be the constant identifier used to track and maintain the provenance of the file. The UID may end up being the same as the file name, but whatever the case be sure to embed the UID inside the file in an appropriate and documented place.

2. **Do not overthink**
   Whether the filename is a randomly generated value or not, be systematic. Think, "Is this logical? Can I spell out the rules easily enough to do batch renaming?" In trying to create the perfectly contained and expressed filename or UID structure, there is a strong temptation to overthink them to the point that they become non-systematic or too idiosyncratic to be logically parsed. If a naming structure is not systematic enough to have a piece of software perform a series of logical renaming steps, there will be lots of manual hours spent retyping names if a mass renaming of files is required at some point in the future.

3. **Do not use filenames as database records**
   Filenames are not the place to cram in a bunch of descriptive and structural information. That's what databases are for! All we require from a filename and ID is that they act as a link to the database record for that unique object. Trying to cram excessive descriptive information into a filename creates unwieldy names and is often futile because of how often conditions or conventions change and new scenarios come up over time. Having filenames that are tied to closely to specific scenarios creates inflexible structures that require non-systematic revision when situations change, which and it puts you in the predicament described in tip #2.

4. **Do not make it machine-unreadable**
   There is often an urge to make a file naming structure decodable by humans, but it also needs to be decodable by computers. Avoid characters that are not URL compatible, that require escape characters, or are reserved by operating systems. Limit options to numbers, letters, periods, and underscores.

5. **Do not assume you will be the first person naming the file**
   When establishing file naming conventions for a collection, most people are considering it in terms of newly derived files reformatted from other sources. In reality, there will be more and more born digital content deposited with archives that already have filenames. In some cases, these can be renamed to fit the archive's naming structure with no loss of information, but at other times, such as with P2 files, the inherited naming structure refers to complex file and directory structures that must be maintained in order to preserve the whole content. Naming structures should be flexible enough to recreate any necessary naming conventions.