Suggestions for file naming conventions in CEL

Why use naming conventions?

- Keeps previous versions from being saved over or lost.
- Visually sorts documents in a folder in a logical manner.
- Makes them easier to find later – especially by other/newer employees.
- Easier to know what information is available – helps for cross-team collaboration.

Good practices in naming conventions

A naming convention is best when the quantity and quality of the agreed upon convention are descriptive enough to make retrieval and search easy by understanding how naming is used. There needs to be a balance, however – including too much information is unwieldly, and makes it difficult to use and understand the naming convention

1. Avoid extra-long folder names and complex hierarchical structures but use information-rich filenames instead.
Think broad and shallow as opposed to narrow and deep:

- Dos:
    - I:\Prod\QA\AssL7_WO_Suzuki_L3688_20090725.xls*
    - I:\Pubs\ Article_eXadox_ File-Naming-Conventions_V03.doc
- Don'ts:
    - I:\Production \Quality Control\Assembly Line7\Work Orders\Clients\Suzuki Motors\ LOT3688_July-25-2009.xls
    - I:\Publications\Articles\eXadox\File-Naming-Conventions_V03.doc

Reason: Complicated folder structures require extra burrowing down when saving and when finding them. By having information in the file name itself, both the search for and the ID of the file is consistent and specific. Note that some systems have a character limitation of 256 characters for the entire path name (Drive letter, folder(s) and file name, so see #2).

2. Abbreviate the content of elements whenever possible.

- Dos:
    - RevQC _QST_2009-Q2.xls
    - MCIM_27643_POD.doc
- Don'ts:
    - Minister of Revenue Quebec _Quebec-Sales-Tax_2009-2ndQuarter.xls
    - MultiCIM-Technologies-Inc_27643_Proof-Of-Delivery.pdf

Reason: Abbreviating helps create concise file names that are easier to read and recognize. Make sure to maintain an easily accessible list of abbreviations.

3. Put sufficient elements in the structure for easy retrieval and identification but do not overdo it.

- Dos:
    - NOVALEC_37507_INVOICE_20090703.pdf
    - FUJITSU_S1500_SPEC_Scanner.pdf
- Don'ts:

- o NOVALEC _INVOICE.pdf
- o FUJITSU_S1500_SPEC_Black _Desktop_Scanner_ ModelReplacesS510_.pdf

Reason: We want to avoid ambiguous search results but too much information adds no returns at retrieval time. The rule of file naming in the olden days was '8.3' – eight characters in the name, and three for the file extension. That's outdated of course, but a reminder to be concise where possible.

4. Use the underscore (_) as a delimiter. Do not use spaces or other characters such as: ! # $ % & ' @ ^ ` ~ + , . ; =) (. Save the hyphen for situations where it is more appropriate, such as dates. Consider using Camel Case (ex. FileNaming.docx).

- Dos:
  - o SMITH-J_AXA_7654-6_POLICY_20120915.pdf
  - o FUJITSU_S1500_SPEC_Scanner.pdf
  - o ThisIsAnExampleOfCamelCase.pdf
- Don'ts:
  - o SMITH-J AXA 7654-6 POLICY 20120915.pdf
  - o FUJITSU $S1500$ SPEC$Scanner.pdf

Reason: The underscore (_) is a quasi-standard for field delimiting. Some search tools do not work with spaces and should be especially avoided for internet files. These characters are also used for other functions in some systems.

On a daily basis, we use Linux, Unix (underlies the Mac OS), Mac OS, iOS and 3 versions of Windows. Making files compatible with all systems helps with collaboration and retrieval.

5. Elements should be ordered from general to specific detail of importance as much as possible.

- Dos:
  - o FY2009_Acme-Corp _Q3_TrialBal_20091015_V02.xls
  - o Production_Paint-Shop_WorkOrder_775-2.xls
- Don'ts:
  - o TrialBal _Q3_20091015_Acme-Corp_V02_FY2009.xls
  - o Paint-Shop_775-2_WorkOrder_Production.xls

Reason: In general, the elements should be ordered logically, in the same sequence that you would normally search for a targeted file.

6. The order of importance rule holds true when elements include date and time stamps.

Dates should be ordered: YEAR, MONTH, DAY. (e.g. YYYYMMDD, YYYYMMDD, YYYYMM). Time should be ordered: HOUR, MINUTES, SECONDS (HHMMSS).

- Dos:
  - o RFQ375_Cables-Unlimited _BID_20091015-1655.pdf
  - o 2009-11-20_AMATProj_Phase1_Report.doc
- Don'ts:
  - o RFQ375_Cables-Unlimited _BID_10152009-1655.pdf
  - o Nov-20-2009_AMATProj_Phase1_Report.doc

Suggestions for file naming conventions in CEL

Reason: To ensure that files are sorted in proper chronological order the most significant date and time components should appear first followed with the least significant components.

7. Personal names within an element should have last name first followed by first names or initials.
Note: Files should generally contain a name because the file pertains SPECIFICALLY to that person (ER, Payroll change, employee review, etc.). NOT because the person is the author or the Project Director.

- Dos:
    - Tate-Peter_SunLife _1-7566-2_POLICY_10YrTerm.pdf
    - SmithJ_ ID3567_ADMIN_WageReview.xls
- Don'ts:
    - Peter-Tate_SunLife _1-7566-2_POLICY_10YearTerm.pdf
    - JSmith_ ID3567_ADMIN_WageReview.xls

Reason: The family name is the standard reference for retrieving records. Having the last name first will ensure that files are sorted in proper alphabetical order.

8. An element for version control should start with V followed by at least 2 digits and should be placed as the last most element. To distinguish between working drafts (i.e. minor revisions) use Vx-01->Vx-99 range and for final draft (i.e. major version release) use V1-00-> V9-xx. (where x =0-9)

- Dos:
    - MCIM_Proposal_V09.doc
    - eXadox_UserManual_V1-02.doc
- Don'ts:
    - MCIM_Proposal_9.doc
    - eXadox_UserManual_V2FinalDraft.doc

Reason: The "V" helps denote that the element pertains to a version number. A minimum of 2 digits with a leading zero is required to ensure that search results are properly sorted. The intent is to avoid the situation where for example, a filename with a "V1-13" will wrongly appear before an identical filename with a "V1-2" version number when sorted in ascending alphabetical/numerical order. To distinguish between working, review and final draft a single digit prefix followed by hyphen "-" is preferred to facilitate proper sorting; using words in the file name such Final, Draft or Review in the filename affect the order and should be avoided.

9. Avoid naming files with extra characters to change sort order.

Reason: This naming convention makes sense only to the author. If there are too many files within a folder to be able to see things without resorting to this, the folder structure probably needs to be simplified – move archival materials to the archives, and create additional subfolders or otherwise move less important folders so that all important folders/files show on the initial screen.

Getting on the same page

A controlled vocabulary, is a definitive list of terms to be used in indexing (human or automated). A controlled vocabulary for a project might actually include multiple authority files for different kinds of terms.

Suggestions for file naming conventions in CEL

Controlled vocabularies are used to ensure consistent indexing, particularly when indexing multiple documents, periodical articles, web pages or sites, etc. They may also be used when indexing a single work, such as an encyclopedia, by multiple indexers.

Controlled vocabularies do not necessarily have any structure or relationships between terms within the list. Controlled vocabularies are often used for name authorities (proper nouns), such as persons, organization names, company names, etc.

Examples:

Defining 'TE' as the standardized abbreviation for 'teaching effectiveness,' (NOT TPEP), or 'ops' for 'operations.'

Sharing the news

It is imperative to share the new guidelines with the organization broadly. It will take time for people to get used to the new paradigm. Not only will you want to document the controlled vocabulary, you will want an SOP or policy around the guidelines.

*File name examples are mostly from the Santaguida cited below.

Bibliography

Brandt, S. (2016, April 25). Library Guides: Data Management for Undergraduate Researchers: Introduction. Retrieved December 19, 2016, from http://guides.lib.purdue.edu/c.php?g=353013&p=2378293

Santaguida, V. (2010, January). Folder and File Naming Convention 10 Rules for Best Practice. Retrieved December 19, 2016, from http://www.exadox.com/en/articles/file-naming-convention-ten-rules-best-practice

Stanford Library. (n.d.). Best practices for file naming. Retrieved December 19, 2016, from https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming

Villella, K. (2011, November 01). Make your point » Naming Conventions/File Architecture. Retrieved December 19, 2016, from http://www.powerframeworks.com/makeyourpoint/?p=413

Taxonomies & Controlled Vocabularies SIG. (n.d.). Retrieved December 19, 2016, from http://www.taxonomies-sig.org/about.htm