File/Folder Naming Problem?

A free-for-all mentality is quite prevalent when it comes to the naming of electronic files and folders. The problem stems from the fact that there are no official standards or file naming and storage policies within most organizations. In such instances, management most likely has not taken the initiative to establish folder and file naming conventions that staff members could follow. Most employees will do what they think is best, but without the proper policies in place the results could be unpredictable and extremely costly. In the days of managing paper files and folders, best practices were in place to ensure that paper-based assets were properly filed to facilitate retrieval. So why not the same for electronic information assets?

Naming Convention - How to?

The strength of a folder and file naming convention is dependent on the proposed naming structure and the quality and quantity of the data elements chosen to build it. It should be of no surprise that for any business activity there is always an ideal naming structure. However, any structured naming convention that attempts to be all encompassing may result in overkill and unwieldiness.

Here are <u>ten basic rules</u> that could serve as a general guideline in structuring folder and file naming conventions:

1. Avoid extra long folder names and complex hierarchical structures but use information-rich filenames instead.
   Dos: Z:\Prod\QA\AssL7_WO_Suzuki_L3688_20090725.xls **||** Z:\Pubs\ Article_eXadox_ File-Naming-Conventions_V03.doc
   Don'ts: Z:\Production \Quality Control\Assembly Line7\Work Orders\Clients\Suzuki Motors\ LOT3688_July-25-2009.xls   **||** Z:\Publications\Articles\eXadox\File-Naming-Conventions_V03.doc
   Reason: Complex hierarchical folder structures require extra browsing at time of storage and at the time of file retrieval. By having all the essential information concisely in the file name itself, both the search and identification of the file is streamlined and more precise.

2. 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: NOVALEC _INVOICE.pdf **||** FUJITSU_S1500_SPEC_Black _Desktop_Scanner_ ModelReplacesS510_.pdf
   Reason: Precision targeted retrieval requires sufficient elements to avoid ambiguous search results but too much information adds undue effort at file naming time with little or no returns at retrieval time.

3. Use the underscore (_) as element delimiter. Do not use spaces or other characters such as: ! # $ % & ' @ ^ ` ~ + , . ; =) (
   Dos: SMITH-J_AXA_7654-6_POLICY_20120915.pdf **||** FUJITSU_S1500_SPEC_Scanner.pdf
   Don'ts: SMITH-J AXA 7654-6 POLICY 20120915.pdf **||** FUJITSU $S1500$ SPEC$Scanner.pdf
   Reason: The underscore (_) is a quasi standard for field delimiting and is the most visually ergonomic character. Some search tools do not work with spaces and should be especially avoided for internet files. Other characters may be interesting but visually confusing and awkward.

4. Use the hyphen (-) to delimit words within an element or capitalize the first letter of each word within an element.
   Dos: Smith-John_AIG _7654-6_POLICY_2009-09-15.pdf **||** WhitePaper_ StructuredFileNamingStrategy.doc
   Don'ts: Smith John AIG 7654 6 POLICY 2009 09 15.pdf **||** White Paper Structured file naming strategy.doc
   Reason: Spaces are poor visual delimiters and some search tools do not work with spaces. The hyphen (-) is a common word delimiter. Alternatively, capitalizing the words within an element is an efficient method of differentiating words but is harder to read .

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

Dos: FY2009_Acme-Corp _Q3_TrialBal_20091015_V02.xls **||** Production_Paint-Shop_WorkOrder_775-2.xls

Don'ts: TrialBal _Q3_20091015_Acme-Corp-V02_FY2009.xls **||** 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: RFQ375_Cables-Unlimited _BID_20091015-1655.pdf **||** 2009-11-20_AMATProj_Phase1_Report.doc

Don'ts: RFQ375_Cables-Unlimited _BID_10152009-1655.pdf **||** Nov-20-2009_AMATProj_Phase1_Report.doc

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 family name first followed by first names or initials.

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 family name first will ensure that files are sorted in proper alphabetical order.

8. 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.

9. 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.

10. Prefix the names of the pertinent sub-folders to the file name of files that are being shared via email or portable storage devices.

Dos: Prod_PS_AssL7_WO_Suzuki_J3688-20090725.xls **||** FY2009_Acme-Corp _Q3_TrialBal_20091015_V02.xls

Don'ts: WO_Suzuki_J3688-20090725.xls **||** Q3_TrialBal_20091015_V02.xls

Reason: Attached files and files shared through portable devices include only the file name and can be totally devoid of the context that is generally provided by the folder structure of origin. To compensate and avoid confusion it is sometimes essential to prefix the name of the subfolder(s) to such file names.

Naming Convention - Solution Planning

Any effort to institute file / folder naming and storage policies may be initiated by anyone in an organization; however, to ensure success it is extremely important that such policies be dictated by the stakeholder with the highest authority. This individual must appreciate and instil to all participants that electronic files:

- are <u>information assets</u> that are the property of the organization.
- are <u>shareable resources</u> although sometimes for restricted users only.
- are <u>critical components</u> to operations that affect productivity and profitability.

Ideally, it would be desirable to achieve consensus on folder and file naming conventions within a team or department; however, this may not always be possible because each participant may believe their ideas are the best. By referring to the above <u>ten basic rules</u> it may possible to resolve the majority of disputes and facilitate the task. If this is not enough it may be necessary to get outside assistance or just simply agree to disagree on any minor issues but ultimately there must be but one standard for each unique business activity whether universally accepted or not.

Execution

Once the file naming and storage policies are finalized, they should be documented and made available to all participants to follow. Everyone should be sensitized to the importance of diligently applying the naming policies and that deviating or compromising in policy adherence could have costly consequences to the organization. Therefore it is incumbent on management that policy adherence be monitored as much as possible because there is always a natural tendency for participants to stray away mostly because of human factors. If this becomes problematic, it is important to be aware that semi-automated structured file naming and storage software exists that can be used to help enforce and monitor adherence to naming and storage policies. For details please refer to *eXadox White Paper: Semi-Automated Structured File Naming and Storage* or visit www.exadox.com.

About the author: Vincent Santaguida is the president of MultiCIM Technologies Inc., a software engineering and business process automation company. He was the conceptual design architect for eXadox, an application that exploits a semi-automated structured file naming and storage strategy as a simple means of organizing and retrieving electronic files and folders.