



File and Record Naming Norms

Bottom Line

This file naming norm draws upon international (ISO) standards, Internet best practices, Connect Care attachment guides, and relevant Alberta Health Services (AHS) Health Information Management procedures. Compliance with the norm can improve the usefulness of shared digital file collections for all stakeholders.

Short but meaningful and relevant

The file or record name should briefly reflect the topic of the information object, as well as the type of information object and the program or initiative it is associated with, all in 60 characters or less.
→ These components need to be ordered in a way that results in easily skimmed lists of files/records that sort to keep like items adjacent.

Delimit without spaces

File names often become part of longer Internet location references (URLs). These links can break with word-wrapping if file names contain spaces. Some browsers substitute codes for spaces that may not be visually recognizable.
→ Separate file name components with an underscore (“_”), use title-style capitalization (first letter) within components and separate words within components (e.g., topic) with a dash (“-”).

Avoid special characters

Avoid “~”, “@”, “#”, “%”, “&”, “?”, “{”, “}”, “<”, “>”, “[” and “]” as these have significance for Internet addresses, programming languages, hypertext and other markup languages. Avoid “\” and “/” as these are used to indicate the relationship between files and directories. Avoid “:”, “.”, “.”, “,” and “*” as these are used to indicate series, truncation or demarcation between file name and file extension.
→ Stick to alphanumeric characters.

Leading zeros

When a number is used in a file name component (such as a version number), like files may not co-sort if they do not have the same number of digits for that component.
→ Always use leading zeros sufficient for the anticipated number range.

ISO date format

Avoid date references in file names as this information is embedded in the meta-data for the file. If a date is essential, place file name components in a consistent order and use ISO standard formats to facilitate co-placement of like files in sorted lists.
→ “YYYYMMDD” is preferred but “YYYY-MM-DD” is acceptable. The year must be 4 digits and the month and day each 2 digits (including a leading zero as needed).

Left to right

Use a standardized hierarchy of file name components, proceeding from left (general) to right (specific): `ORG_PROG_TYPE_TOPIC_VER_INITIALS.EXT`.

Versioning is not typically needed (as this is embedded in meta-data) but can be indicated by number or date (ISO format), possibly followed by last editor(s) initials.

→ Indicate the organization (ORG) and initiative (PROG) if multiple groups may produce the same kind of document on similar topics. A category (TYPE) of document can help to group similar objects in long lists. The title or content (TOPIC) should be briefly indicated with a few words separated by dashes.

Examples

AHS_CC_TIP_Business-Continuity.doc

AHS_CC_NORM_File-Naming_CMIO.docx



Objectives

The intent of this norm is to provide direction for AHS Connect Care clinical information system (CIS) stakeholders about a common approach to the naming of files, folders, records and other artifacts that may be gathered in support of CIS activities and stored or linked on SharePoint, shared drive, website, online manual or other electronic collections. Consistent naming can dramatically improve file finding and sorting, as well as Internet link stability.

Background

Countless documents, policies, briefings, presentations, requests and other information “objects” accrue in association with Connect Care activities. Individual objects may remain “off-grid” while in development and later end up on a shared document repository (e.g., shared network drive) or become part of an official project archive using Microsoft SharePoint. Key objects may, in turn, be linked from user-facing handbooks, manuals, dashboards or other resources.

Given the typical origin of most information objects, file names usually reflect personal naming habits. These rarely express shared standards, and typically reflect word processor default title-based file names. This may be okay for the author, who can find files in personal collections. But when the object appears in shared collections, inconsistent naming can make it difficult for others to recognize what the file is about or to find it among a growing jumble of collected objects.

File naming conventions have evolved from systems that were optimized for index cards to ones better attuned to computer file managers and, most recently, natural language processors. Current best practices build names in ways that are Internet-safe and enterprise-aware. The goal is to make it easier to search and browse large collections via content management systems, online databases and Internet search engines.

Significance

File naming practices affect:

- Speed and accuracy of automated searches for specific files or groups of files.
- Likelihood that Internet versions of file names will remain intact when shared via email and other communications channels.
- Likelihood that Internet links to file locations will remain unambiguous and free of characters that may have unexpected effects in hypertext links.
- Ease of finding files in lists displayed by typical information collection managers.
- Ability to still recognize a file’s likely purpose when the file is separated from its folder hierarchy.
- Grouping of files relating to similar topics or initiatives.

Applicability

This guide should be followed by all who manage documents associated with CMIO or Connect Care CIS digital file collections and online repositories.

Guides

This file naming norm draws upon international (ISO) standards, Internet best practices, Connect Care attachment guides, and relevant AHS Health Information Management procedures. It emphasizes requirements for efficient, unambiguous management of files that may end up in digital repositories or Internet resources. Other conventions are less particular about special characters and word delimiters, as these parts of the norm impact Internet but not paper-based locators.



- **File and record names should be short but meaningful and relevant.**
 - Word processing programs (e.g., Microsoft Word) will generate a default name for a new file using any header or title words found in the document, resulting in inconsistent formatting and verbose titles.
 - The file or record name should briefly reflect the topic of the information object, as well as the type of information object and the program or initiative it is associated with. These components need to be ordered in a way that promotes consistent sorting and filtering of large lists.
 - Avoid superfluous words that are best reflected by the actual document content, such as “draft”, “correspondence”, or “approved”.
 - Avoid redundancy where file name components are already implicit in directory structures if the file will usually appear in organized collections.
 - Limit file names to 60 characters or less. Anything more is difficult to include in typical document references. The combination of a file name and all folder names required to locate the file should not exceed 255 characters.

- **File name components and words within components should not have spaces as delimiters.**
 - File names often become part of longer Internet location references (URLs) that may appear within documents, email or other communications. As communication programs tend to word-wrap paragraphs, if file names contain spaces, their resulting links are at risk for being “broken” across lines. This defeats any automated recognition of Internet locations (for creating of hypertext and links).
 - Although spaces are handled by modern disk operating systems, errors can occur when the file name becomes part of an Internet-based file management system (e.g., SharePoint) or website. These commonly substitute “escape codes” (e.g., “%20”) for spaces.
 - Separate file name components with an underscore (“_”).
 - Programs like SharePoint typically concatenate words with an underscore. Current Internet search engines and file management systems handle underscores well.
 - Separate words within file name components with a dash (“-”).
 - Internet search engines and file management programs treat words separated by dashes as discrete terms.
 - If component words represent discrete concepts that might respond best to a search that combines the concepts with “AND”, then it is best to capitalize each word in the component (separated by dashes). If in doubt, default to title-style capitalization (first letter of each word).

- **Avoid special characters that may be misinterpreted or automatically substituted in an Internet or disk file management context.**
 - Avoid “~”, “@”, “#”, “%”, “&” and “?” as these are used as delimiters or escape codes within Internet addresses.
 - Avoid “\” and “/” as these are used to indicate the relationship between files and directories.
 - Avoid “<”, “>”, “{”, “}”, “[” and “]” as these have special significance for programming languages, hypertext and other markup languages.
 - Avoid single or double quotation marks within file names.
 - Avoid “,”, “:”, “.”, “,” and “*” as these are used to indicate series, truncation or demarcation between file name and file extension.
 - It is safest to stick to alphanumeric characters.



- **When a number is used in a file name component (such as a date or version number), always use leading zeros so that the number range used by like files has the same number of digits for that component.**
 - If a numeric indicator has three digits, then version 3 would be “003” and version 300 would be “300”.
- **Follow international (ISO) standards for indicating dates.**
 - Dates typically do not need to be included in file names. Links to resources from Internet pages need to be stable, and so file names should not vary by date.
 - The date of first authorship and last edit are already embedded in the meta-data for most files, and so are not needed in file names.
 - If a date must be used in the actual file name, “YYYYMMDD” is preferred but “YYYY-MM-DD” is acceptable.
 - The year must be 4 digits, and the month and day each 2 digits (including a leading zero as needed).
 - If a specific day is not indicated, use “YYYYMM” or “YYYY-MM”.
- **Use a standardized hierarchy of file name components, proceeding from left (general) to right (specific).**
 - `ORG_PROG_TYPE_Topic.EXT`
 - `ORG` = (optional) Indication of organization, group or general sponsorship of the file.
 - `PROG` = Brief indicator of the department, program, or initiative that the file is associated with.
 - `TYPE` = (optional) May be reflected by short indicators of different classes of documents where there is a common understanding of what these are:
 - “BN” – briefing note
 - “BYTE” – short conceptual introduction
 - “DR” – decision request
 - “FAQ” – frequently asked questions
 - “GUIDE” – overview of a workflow
 - “POLICY” – policy or governance document
 - “PROCESS” – procedure or process document
 - “PROTOCOL” – standard or protocol
 - “GUIDE” – guideline
 - “TIP” – short explanation about how to do something
 - “REPORT”, “CHARTER”, “PROPOSAL” or “GRANT” may also be appropriate
 - `Topic` = Brief indication of the topic of the information object, using dashes to separate words. If the topic is a person (e.g., personnel records), always place the last name first, followed by the first name and, if necessary, any initial(s), separating all with dashes.
 - `VER` and `INITIALS` = Version and editor initials should not be used in final-form documents but may be appended for documents in development, separated by underscores. Use ISO date format.
 - `EXT` = Usually indicates the type or identity of the software required to open, display and/or edit the file. File name extensions are usually assigned automatically by the originating software and should not be changed.
- **When choosing standardized sets of organization, program, type and topic names, consider how the intended users might search for files using specific keywords or combinations of letters.**



Example

The following example illustrates application of the above principles.

File Name	Comment
AHS_CC_TIP_Business-Continuity_CMIO.doc	The associated tip sheet (Type="TIP") has an Alberta Health Services ("AHS") scope within its Connect Care ("CC") initiative. Multiple versions of the TIP exist, with this one specifically formatted to address the needs of prescribers ("CMIO"). Internet searches for the concepts "business" and "continuity" will rank this file-find higher.

Definitions

(See glossary.connect-care.ca)

- **Record**
Documents, data or information of any kind and in any format (e.g., documents, spreadsheets, databases, emails, blogs, wikis and website pages) created, received, recorded and maintained by AHS as part of its services or business.
- **URL**
Universal Resource Locator, a standardized way to refer to the Internet location and identity of a specific page, object or document on the World Wide Web.

Resources

- Creating good file names for websites and blogs
 - <http://www.thesitewizard.com/webdesign/create-good-filenames.shtml>
- Simple naming conventions that improve end user experience in SharePoint sites
 - <http://www.networkworld.com/article/2221731/microsoft-subnet/simple-naming-conventions-that-improve-end-user-experience-in-sharepoint-sites.html>
- Cross-platform filename best practices and conventions
 - <https://support.apple.com/en-us/HT202808>
- SharePoint tips – Naming convention
 - <https://www.linkedin.com/pulse/sharepoint-tips-naming-convention-nick-rich>
- Best practices for file naming
 - <https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming>
- Folder and file naming convention – 10 rules for best practice
 - <http://www.exadox.com/en/articles/file-naming-convention-ten-rules-best-practice>