ELECTRONIC RECORDS MANAGEMENT GUIDELINES

FILE NAMING

Summary

A file name is the chief identifier for a record. In the world of electronic records, the record's file name provides metadata that place the record in context with other records, records series, and records retention schedules. In most organizations, the policy for naming a file (and hence a record) is left to individuals or to groups of individuals (e.g., departments, committees). Consider establishing an agency-wide file naming policy that complements your electronic records management strategy.

Consistently named records foster collaboration based on mutual understanding of how to name files and use file names (including the file name metadata). Consistently named records also help you to meet your legal requirements. Legally, your records must be trustworthy, complete, accessible, admissible in court, and durable for as long as your approved records retention schedules require. Records that are consistently and logically named are easier to manage to meet these requirements.

When each staff member consistently names electronic records, other staff members will be able to look at a record's file name and use that information to recognize the contents and characteristics of the record and make decisions about it. For example, a staff member could see that "HR0035broch96_97P.pdf" is a brochure about a House bill (HR0035) in the 1996-1997 session that is available to the public.

Legal Framework

For more information on the legal framework you must consider when naming files, refer to the chapter *Records Management in an Electronic Environment* in the Electronic Records Management Guidelines and Appendix A6 of the *Trustworthy Information Systems Handbook*. Also review the requirements of the:

- Public Records Act [PRA] (*Code of Laws of South Carolina, 1976*, Section 30-1-10 through 30-1-140, as amended) available at www.scstatehouse.org/code/t30c001.htm, which supports government accountability by mandating the use of retention schedules to manage records of South Carolina public entities. This law governs the management of all records created by agencies or entities supported in whole or in part by public funds in South Carolina. Section 30-1-70 establishes your responsibility to protect the records you create and to make them available for easy use. The act does not discriminate between media types. Therefore, records created or formatted electronically are covered under the act.

Differences Among File Names, File Paths, and Addresses

A *file name* is the name of the file as it stands alone. "CommitteeAMinutes021401.doc" is an example of a file name.

The *file path* shows the location of the file on your network or local computer. For example, the file "CommitteeAMinutes021401.doc" might be stored in a series of nested directories for all committees as:
X:\Committees\CommitteeA\Minutes\2001\February\CommitteeAMinutes021401.doc

An *address* describes the location of a file accessible by a web browser. For example, a map of a public park named "Smith Park" might have the following address: www.parks.org/smith.html

Challenges in File Naming

As you develop your policy, you will encounter the following challenges in file naming:

- *Version control*. You will need to determine how and whether to indicate the version of the record. Some organizations put current and obsolete drafts in different electronic file folders without altering the file name. However, when these records are moved from the active electronic file folder to another storage area, identical file names may conflict and cause confusion.
- *Uniqueness*. To avoid file names conflicting when they are moved from one location to another, each record's file name should be unique and independent from its location. For example, if

MORE →

letters are simply named with the word *letter* and the date, [Letter1May2005.doc] they are not independent from location because they could fit into any records series that contains letters, and all letters sent on that date would have the same file name.

- *Persistence over time*. File names should outlast the records creator who originally named the file. With good stakeholder and staff input, and training, your staff should be able to develop file names that make sense to others without needing to rely on the file creators for interpretation.
- *Access and ease of use*. The file-naming policy should be simple and straightforward. A simple policy will help staff members logically and easily name records and help ensure that records are accessible to other staff members and/or to the public. A simple policy that is used regularly results in records that are consistently named and easier to organize and access.
- *Ease of administration*. The policy should work with your computer infrastructure, so that you can monitor policy compliance, manage records and records series, gather metadata, and perform other administrative tasks easily and in compliance with all legal requirements. For example, if all the records in a specific records series are easily identifiable by file name, they will be easier to gather and manage.
- *Scalability*. Consider how scalable your file naming policy needs to be. For example, if you want to include the project number, don't limit your project numbers to two digits, or you can only have ninety-nine projects.

Common File Name Elements

When developing your file naming policy, you may wish to include some of the following common elements:

- Version number (e.g., version 1 [v1, vers1])
- Date of creation (e.g., February 14, 2005 [021405, 02_14_05, 20050214])
- Name of creator (e.g., Rupert B. Smith [RBSmith, RBS])
- Description of content (e.g., media kit [medkit, mk])
- Name of intended audience (e.g., general public [pub])
- Name of group associated with the record (e.g., Committee ABC [CommABC])
- Release date (e.g., released on June 11, 2001 at 8:00 a.m. central time [061101_0800CT])
- Publication date (e.g., published on December 24, 2003 [pub122403])
- Project number (e.g., project number 739 [PN739])
- Department number (e.g., Department 140 [Dept140])
- Records series (e.g., SeriesX)

Internet File Naming Protocols

Several file naming protocols are currently in use on the Internet. They all fall under the category of Uniform Resource Identifiers (URIs). URIs are short text strings that identify Internet resources (e.g., documents, images, electronic mailboxes). These text strings commonly appear in the address bar of a web browser and are referred to as an Internet addresses [e.g., http://www.state.sc.us/scdah/]. The first part of a URI specifies the *transfer protocol* in use (the method for transmitting the file from one device to another, such as hypertext transfer protocol [e.g., http://]. The second part specifies the address, often including the domain name, of the file [e.g., www.state.sc.us/scdah/].

Within the broad grouping of URIs are:

- *Uniform Resource Locators (URLs)*. URLs are specific schemes that allow browsers and other software to access resources on the Internet. URLs indicate the resource's location [e.g., address and name].
- *Persistent Uniform Resource Locators (PURLs)*. PURLs are functionally URLs, but act as an intermediary for the URL of a web site by redirecting the browser to a PURL server instead of the server that hosts the web site. The PURL server associates the PURL with the real URL and returns the web page to the viewer's browser. [http://purl.oclc.org/]
- *Uniform Resource Names (URNs)*. URNs are designed to serve as persistent, location-independent resource names. URNs are intended to overcome the problem of persistence and location-independence by providing a long-term identifier for resources. URNs use a resolution service to enable a web browser to use the URN to find the URL location for the resource. An example of a URN is - isbn:0262194643

MORE →

Domain Names

Common practice is to include the domain name in the URI. A domain name, such as “microsoft.com,” is nearly always a part of the URL, because URLs identify resources by location. Review the resources in the Annotated List of Resources for more information on domain names. Domain names are administered by the Domain Name System (DNS).

URL Protocols

You will encounter several common types of URL transfer protocols, including:

- *Hypertext Transfer Protocol (HTTP)*. This is the most common type of URL protocol accessed on the Internet [e.g., http://www.state.sc.us/scdah/].
- *File Transfer Protocol (FTP)*. This protocol type is commonly used to transfer large files via the Internet [e.g., ftp://ftp.rootsweb.com/pub/usgenweb/sc/sccolony.txt].
- *Gopher*. This protocol was used primarily in academic and governmental settings, and is rarely used today.
- *Telnet*. This protocol allows users to control the activity on another computer and participate in interactive sites for such activities as games, live chats, and text information exchange. [e.g., telnet://leo.scsl.state.sc.us:23]
- *Mailto*. This protocol is for e-mail exchange. [e.g., mailto:generalinfo@scprt.com]

Challenges in Internet-Based File Naming

Naming for the Internet is particularly challenging. File names should meet your general criteria, especially for uniqueness, independence from location, and persistence over time. The file names should persist even if you move the files to another server, reorganize your web site, or use another software program or method for producing your web pages.

A carefully constructed policy for naming Internet-delivered files will ensure that:

- *Your web site links stay live*. Links contain embedded information about the location of the resource being linked to. Moving files from one server to another may result in dead links. If you develop a policy that builds in persistence and location-independence, you should be able to avoid this problem.
- *You can more easily manage your web site records*. Because the file names are independent of location, you can be assured that you will be able to find records if they are reorganized. For example, if a department within your agency reorganizes its web pages and moves some files to another server, as long as the file names of the records are independent of location, you can still efficiently manage and archive them.
- *Your Internet-accessed files mesh with your other electronic files*. By integrating your file naming policy with that used for other electronic records, your public records will remain accessible as long as necessary and confidential information will be protected as appropriate.

Six steps to better file naming

1. *Establish naming conventions* — Form a coordinating committee that includes records managers, information technology administrators, executive management and users to establish a consistent method for naming files in your agency. In addition to the other benefits, establishing file naming rules helps meet legal requirements by making records much more accessible.
2. *Determine file naming boundaries* — Pay close attention to the freedom you give staff members (and outside vendors) in naming files. Provide guidelines and training on file naming. One individual will not be able to manage file naming for every electronic record, so you will need to rely on staff members and vendors to name files in compliance with your policy. By providing guidelines and training, you can maximize policy compliance in a way that meets your operational and legal requirements. New employees should receive training as part of your agency’s orientation process. All employees should receive follow up training at established intervals to encourage good habits and maintain consistency.
3. *Identify the record or official copy* — You may have numerous copies of records in your office stored on individual hard drives and network servers or moving between you and your clients. The *official* copy is the record that proves your agency performed a specific function correctly. You must maintain the official copy according to the retention period established in your retention schedule. The *original* is the document sent to the customer, or the one used to initiate an action. If you send a letter, memo or report, the original goes to the recipient and you file a copy as the official copy; When you receive a letter, memo or

MORE →

report the reverse is true — you file the original as the official copy. It is very easy to *copy* documents in an electronic environment. Procedures should be set out by your records manager and coordinating committee that define which copy is the official copy and that specify its location, content and access controls. Official copies should be kept in a secure, trustworthy system that provides audit trails. Copies should be destroyed when no longer needed for reference and should never be kept longer than the official copy.

4. *Determine the relationship to and connection with paper records* — Determine how the names of your electronic records relate to the names of paper files you have stored. Because electronic records may be part of records series that include paper records, the file naming policy for electronic records should fit logically with your paper records naming. For example, a letter published on a web site might be part of a records series that includes additional paper documents in a file folder. By ensuring that the file names are consistent, you can more easily manage the records series. In order to control records of different media you could establish an index that link the records together.
5. *Create an index* — The more your agency depends on electronic systems to store and keep information the more difficult it becomes to find. Creating and maintaining an up to date index of your directories on the root directory can help. The directories should be listed with a brief description of its contents and the naming convention applied to it.
6. *Be consistent!* — More than anything else, consistent practices will help you avoid pitfalls associated with poor records management. Naming records consistently across your organization will make finding and managing information easier.

Annotated List of Resources

Primary Resources

Cool URIs Don't Change. In: *Style Guide for Online Hypertext*. Cambridge, MA: World Wide Web Consortium (W3C), 1998. www.w3.org/Provider/Style/URI

This section of the complete style guide discusses the file naming concepts for the World Wide Web to ensure the accuracy of links and the longevity of the names.

Naming and Addressing: URIs, URLs,.... www.w3.org/Addressing

These web pages describe the relationship of URIs, URLs, and URNs. The pages also provide links and other information about other file naming topics for the web, such as metadata, markup languages, events, and history.

PURL. purl.oclc.org/

The OCLC PURL Service provides a comprehensive introduction to the subject of PURLs. Available from this web site are Frequently Asked Questions on PURLs, introductions to the subject, and the opportunity to create and modify a PURL.

Additional Resources

Identifiers for Digital Resources. Washington, D.C.: Library of Congress, National Digital Library Program, 1996. memory.loc.gov/ammem/award/docs/identifiers.html

These web pages describe the desirable characteristics for file naming for digital records. For illustrative purposes, the pages use the American Memory Collection as a case study for a file naming scheme.

Mims, J. "Files Control." In *Records Management: A Practical Guide*. Washington, D.C.: International City County Management Agency, 1996: 73-84.

This chapter on file management discusses such topics as filing systems, filing system creation, filing system maintenance, and filing system equipment. It also offers information on troubleshooting file system control. The content focuses primarily on paper systems, but the management principles apply across all media.

South Carolina Department of Archives and History. *Trustworthy Information Systems Handbook*. Version 1, July 2004. www.state.sc.us/scdah/erg/tis.htm

This handbook provides an overview for all stakeholders involved in government electronic records management. Topics center around ensuring accountability to elected officials and citizens by developing systems that create reliable and authentic information and records. The handbook outlines the characteristics that define trustworthy information, offers a methodology for ensuring trustworthiness, and provides a series of worksheets and tools for evaluating and refining system design and documentation.