Document management system

From Wikipedia, the free encyclopedia

This is the current revision of this page, as edited by Mindmatrix (talk | contribs) at 14:17, 27 April 2010. The present address (URL) is a permanent link to this version.

(diff) ← Previous revision | Current revision (diff) | Newer revision → (diff)

A **document management system** (DMS) is a computer system (or set of computer programs) used to track and store electronic documents and/or images of paper documents. The term has some overlap with the concepts of content management systems. It is often viewed as a component of enterprise content management (ECM) systems and related to digital asset management, document imaging, workflow systems and records management systems.

Contents

Overview

In the broadest sense, document management systems can range from a shoebox all the way to an enterprise content management system. There are several common issues that are involved in managing documents, whether the system is an informal, ad-hoc, paper-based method for one person or if it is a formal, structured, computer-enhanced system for many people across multiple offices. Most methods for managing documents address the following areas:

Location	Where will documents be stored? Where will people need to go to access documents? Physical journeys to filing cabinets and file rooms are analogous to the onscreen navigation required to use a document management system.
Filing	How will documents be filed? What methods will be used to organize or index the documents to assist in later retrieval? Document management systems will typically use a database to store metadata about documents and a File System to store the actual physical files.
Retrieval	How will documents be found? Typically, retrieval encompasses both browsing through documents and searching for specific information. What kinds of document information are indexed for rapid retrieval?

Security	How will documents be kept secure? How will unauthorized personnel be prevented from reading, modifying or destroying documents?
Disaster recovery	How can documents be recovered in case of destruction from fires, floods or natural disasters?
Retention period	How long should documents be kept, i.e. retained? As organizations grow and regulations increase, informal guidelines for keeping various types of documents give way to more formal records management practices.
Archiving	How can documents be preserved for future readability?
Distribution	How can documents be available to the people that need them?
Workflow	If documents need to pass from one person to another, what are the rules for how their work should flow?
Creation	How are documents created? This question becomes important when multiple people need to collaborate, and the logistics of version control and authoring arise.
Authenticity	Is there a way to vouch for the authenticity of a document ?
Traceability	When, where and by whom are documents created, modified, published and stored [1]?

History

Beginning in the 1980s, a number of vendors began developing systems to manage paper-based documents. These systems managed paper documents, which included not only printed and published documents, but also photos, prints, etc.

Later, a second style of system was developed, to manage electronic documents, i.e., all those documents, or files, created on computers, and often stored on local user file systems. The earliest electronic document management (EDM) systems were either developed to manage proprietary file types, or a limited number of file formats. Many of these systems were later referred to as document imaging systems, because the main capabilities were capture, storage, indexing and retrieval of image file formats. These systems enabled an organization to capture faxes and forms, save copies of the documents as images, and store the image files in the repository for security and quick retrieval (retrieval was possible because the system handled the extraction of the text from the document as it was captured, and the text indexer provided text retrieval capabilities).

EDM systems evolved to where the system was able to manage any type of file format that could be stored on the network. The applications grew to encompass electronic documents, collaboration tools, security, and auditing capabilities.

Components

Document management systems commonly provide storage, versioning, metadata, security, as well as indexing and retrieval capabilities. Here is a description of these components:

Metadata

> Metadata is typically stored for each document. Metadata may, for example, include the date the document was stored and the identity of the user storing it. The DMS may also extract metadata from the document automatically or prompt the user to add metadata. Some systems also use optical character recognition on scanned images, or perform text extraction on electronic documents. The resulting extracted text can be used to assist users in locating documents by identifying probable keywords or

providing for full text search capability, or can be used on its own. Extracted text can also be stored as a component of metadata, stored with the image, or separately as a source for searching document collections.

Integration

Many document management systems attempt to integrate document management directly into other applications, so that users may retrieve existing documents directly from the document management system repository, make changes, and save the changed document back to the repository as a new version, all without leaving the application. Such integration is commonly available for office suites and e-mail or collaboration/groupware software. Integration often uses open standards such as ODMA, LDAP, WebDAV and SOAP to allow integration with other software and compliance with internal controls.[*citation needed*]

Capture

Images of paper documents using scanners or multifunction printers. Optical character recognition (OCR) software is often used, whether integrated into the hardware or as stand-alone software, in order to convert digital images into machine readable text. Optical mark recognition (OMR) software is sometimes used to extract values of check-boxes or bubbles.

Indexing

Track electronic documents. Indexing may be as simple as keeping track of unique document identifiers; but often it takes a more complex form, providing classification through the documents' metadata or even through word indexes extracted from the documents' contents. Indexing exists mainly to support retrieval. One area of critical importance for rapid retrieval is the creation of an index topology.

Storage

Store electronic documents. Storage of the documents often includes management of those same documents; where they are stored, for how long, migration of the documents from one storage media to another (hierarchical storage management) and eventual document destruction.

Retrieval

Retrieve the electronic documents from the storage. Although the notion of retrieving a particular document is simple, retrieval in the electronic context can be quite complex and powerful. Simple retrieval of individual documents can be supported by allowing the user to specify the unique document identifier, and having the system use the basic index (or a non-indexed query on its data store) to retrieve the document. More flexible retrieval allows the user to specify partial search terms involving the document identifier and/or parts of the expected metadata. This would typically return a list of documents which match the user's search terms. Some systems provide the capability to specify a Boolean expression containing multiple keywords or example phrases expected to exist within the documents' contents. The retrieval for this kind of query may be supported by previously-built indexes, or may perform more time-consuming searches through the documents' contents to return a list of the potentially relevant documents. *See also Document retrieval.*

Distribution

A published document for distribution has to be in a format that can not be easily altered. As a common practice in law regulated industries, an original master copy of the document is usually never used for distribution other than archiving. If a document is to be distributed electronically in a regulatory environment, then the equipment tasking the job has to be quality endorsed AND validated. Similarly quality endorsed electronic distribution carriers have to be used. This approach applies to both of the systems by which the document is to be inter-exchanged, if the integrity of the document is highly in demand.

Security

Document security is vital in many document management applications. Compliance requirements for

certain documents can be quite complex depending on the type of documents. For instance the Health Insurance Portability and Accountability Act (HIPAA) requirements dictate that medical documents have certain security requirements. Some document management systems have a rights management module that allows an administrator to give access to documents based on type to only certain people or groups of people.

Workflow

Workflow is a complex problem and some document management systems have a built in workflow module. There are different types of workflow. Usage depends on the environment the electronic document management system (EDMS) is applied to. Manual workflow requires a user to view the document and decide who to send it to. Rules-based workflow allows an administrator to create a rule that dictates the flow of the document through an organization: for instance, an invoice passes through an approval process and then is routed to the accounts payable department. Dynamic rules allow for branches to be created in a workflow process. A simple example would be to enter an invoice amount and if the amount is lower than a certain set amount, it follows different routes through the organization.

Collaboration

Collaboration should be inherent in an EDMS. In its basic form, a collaborative EDMS should allow documents to be retrieved and worked on by an authorized user. Access should be blocked to other users while work is being performed on the document. Other advanced forms of collaboration allow multiple users to view and modify (or markup) a document at the same time in a collaboration session. The resulting document should be viewable in its final shape, while also storing the markups done by each individual user during the collaboration session.

Versioning

Versioning is a process by which documents are checked in or out of the document management system, allowing users to retrieve previous versions and to continue work from a selected point. Versioning is useful for documents that change over time and require updating, but it may be necessary to go back to or reference a previous copy.

Searching

Finds documents and folders using template attributes or full text search. Documents can be searched using various attributes and document content

Publishing

Publishing a document is sometimes tedious and involves the procedures of proofreading, peer or public reviewing, authorizing, printing and approving etc. Those steps ensure prudence and logic thinking. Any careless handling may result in the inaccuracy of the document and therefore mislead or upset its users and readers. In law regulated industries, some of the procedures have to be completed as evidenced by their corresponding signatures and the date(s) on which the document was signed. Refer to the ISO divisions of ICS 01.140.40 and 35.240.30 for further information.[2][3]

The published document should be in a format that is not easily altered without a specific knowledge or tools, and yet it is read-only or portable.[4]

Standardization

Many industry associations publish their own lists of particular document control standards that are used in their particular field. The following is the list of some of the relevant ISO documents. Divisions ICS 01.140.10 and 01.140.20.[5][6] The ISO has also published a series of standards regarding the technical documentation, covered by the division of 01.110.[7]

- ISO 2709:1996 Information and documentation—Format for information exchange

- ISO 15836:2009 which replaces ISO 15836:2003 Information and documentation — The Dublin Core metadata element set
- ISO 21127:2006 Information and documentation—A reference ontology for the interchange of cultural heritage information
- ISO 23950:1998 Information and documentation—Information retrieval (Z39.50) — Application service definition and protocol specification.
- ISO/CD 10244 Document management—Business process/workflow baselining and analysis associated with EDMS technologies
- ISO 32000 — portable document format

Document taxonomy (classification) and nomenclature

See also

- Content management system
- Data proliferation
- Document automation
- Documentation
- Information repository
- Information science
- Intelligent document
- Library science
- Revision control

References

1. ^ Food and Drug Administration (2007). "Guidance for Industry: Computerized Sysems Used in Clinical Investigations" (http://www.fda.gov/cder/guidance/7359fnl.pdf) . http://www.fda.gov/cder/guidance/7359fnl.pdf. Retrieved 9 April 2009.
2. ^ International Organization for Standardization. "01.140.40: Publishing" (http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=140&ICS3=40&) . http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=140&ICS3=40&. Retrieved 14 July 2008.
3. ^ International Organization for Standardization. "35.240.30: IT applications in information, documentation and publishing" (http://www.iso.org/iso/catalogue_ics_browse?ICS1=35&ICS2=240&ICS3=30&) . http://www.iso.org/iso/catalogue_ics_browse?ICS1=35&ICS2=240&ICS3=30&. Retrieved 14 July 2008.
4. ^ OnSphere Corporation. "SOP Document Management in a Validated Environments" (http://www.onsphere.com/documents/SOPDocumentManagementinValidatedEnvironments.pdf) . http://www.onsphere.com/documents/SOPDocumentManagementinValidatedEnvironments.pdf. Retrieved 4 April 2009.
5. ^ International Organization for Standardization. "01.140.10: Writing and transliteration" (http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=140&ICS3=10&) . http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=140&ICS3=10&. Retrieved 14 July 2008.
6. ^ International Organization for Standardization. "01.140.20: Information sciences" (http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=140&ICS3=20&) . http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=140&ICS3=20&. Retrieved 14 July 2008.
7. ^ International Organization for Standardization. "01.110: Technical product documentation" (http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=110&) . http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=01&ICS2=110&. Retrieved 15 July 2008.

External links

- Software/Document Management (http://www.dmoz.org/Computers/Software/Document_Management/) at the Open Directory Project
- Miles L. Mathieu, Ernest A. Capozzoli (2002) (PDF). *The Paperless Office: Accepting Digitized data* (http://www.cabinetng.com/media/Paperless%20Office_Troy%20State.pdf) . Troy State University. http://www.cabinetng.com/media/Paperless%20Office_Troy%20State.pdf.
- Kevin Craine. *Excerpts from* Designing a Document Strategy (http://www.document-strategy.com/excerpts.html) . Craine Communications Group. http://www.document-strategy.com/excerpts.html.

Retrieved from "http://en.wikipedia.org/wiki/Document_management_system"
Categories: Content management systems | Information technology management | Document Management System

---