Considerations for Content Management Systems

Prepared By Robert Blatt, MIT, LIT,
Chair AIIM Implementation Guidelines Committee
Chair, AIIM C27 Document Management

Table of Contents

1. Introduction

For purposes of discussion, the terms "document management" and "content management" can be considered to be synonymous. As the Electronic Content Management industry (previously referred to as the document management industry) has matured over the years the ability to store electronic information has greatly expanded from hardcopy document scanning into digital images in the early 1980's to the management of any digital or electronic document that today is referred to as electronically stored information.

This document provides detailed information associated with the primary technologies commonly associated with Content or Document Management Systems. Contents of this document have been extracted from the content/document management industry best practice ARP1, *Recommended Practice: Analysis, Selection, and Implementation of Electronic Document Management Systems (EDMS),* produced by the AIIM standards program (www.aiim.org/standards). The AIIM standards program follows strict procedures when developing standards and best practices to ensure complete and accurate up-to-date information is provided. All AIIM documents are prepared by standards committees consisting of experts in the relevant field, end-users, and members of the public ensuring all views and input is received.

"Industry guidelines and standards enable organizations to follow industry accepted practices and procedures. Standards and recommended practices specified in a federal, state or local law or regulations are required specifically in the area covered by the law or regulation. Users wishing to require adherence to a standard or recommended practice should specify them in their procurement documents and contracts since this is the only way a vendor is required to meet a standard. Users of standards should also be careful to specify exactly what requirements in a standard must be met. It is possible for a system to 'meet' a standard and still not deliver the required results if the contract is not specific about the contents of the standard or recommended practice.

Following industry guidelines and standards will further improve the ability of an organization to implement the selected technologies following policies and procedures found necessary, throughout the industry, to implement highly successful systems. These guidelines and standards also enable the organizations to implement products and technologies meeting their specific needs while being able to share information with other organizations who may, or may not, have the same product installed. Industry guidelines provide specific information to users that will enable them to gain information necessary to successfully prepare for, select, and implement the desired technology."[1]

"Terms and acronyms associated with various aspects of EDMS technologies commonly change over time, as technology developers and vendors update product lines and solutions to address customer requirements. In most cases the new terms and acronyms reflect updates and changes to how these technologies are utilized, typically by incorporating additional levels of functionality and very rarely resulting in an entirely new core technology.

This is important to note as the core EDMS technologies are constantly maturing and solution providers are identifying not only new approaches to addressing organizational issues and requirements, but also expanding the use of these technologies into areas previously unconsidered.

It is important to note that as the various technologies associated with storing and managing electronically stored information continue to mature and change, terms and acronyms will continue to change and, at times, be used to denote something different than previously used in the past. As such, organizations are constantly challenged to keep pace with how an updated

---

[1] *ARP1 – 2009, Recommended Practice: Analysis, Selection, and Implementation of Electronic Document Management Systems (EDMS).* Silver Spring, MD: AIIM, 2009. Web (sec. 5.1). Permission for use has been granted by AIIM International. Material attributed to *ARP1 – 2009* and other AIIM publications cited are protected by U.S. Copyright Law (Title 17, U.S. Code).

technology is currently being referenced, especially when the same core technology is referenced differently between vendors, and at times various groups of suppliers."[2]

Because this is a complex, technical subject matter, readers should refer to the Glossary whenever they are uncertain of a term.

2. Assessing Existing ECM Solutions

The industry standard for assessing existing systems was recently published. ANSI/AIIM 25: 2012, *Assessing Trusted Systems for Compliance with Industry Standards and Best Practices* was prepared by the ANSI/AIIM C27.3 Trusted Solution standardization committee.[3]

ANSI/AIIM 25 provides methodology for assessing ECM systems to determine if they are reproducing reliable and accurate renditions of information originally put into the system.

ECM Systems are way more than just a file cabinet for your electronic files. Today almost all of our business documents are created electronically, as well as transmitted, acted upon, edited and stored in an electronic format. In many cases "documents" never reach a physical or paper form – the entire lifecycle is spent in electronic format. ECM systems allow for work flow as well as storing and maintaining records.

Investing in and installing the hardware/software and scanning for your business documents and storing them electronically doesn't automatically mean you have a trusted ECM system. As ANSI/AIIM 25 sets out, a trusted ECM system encompasses all aspects governing the lifecycle of the data from creation or ingestion through destruction and special focus must be placed on the policy and procedures governing the human factor that interacts with the ECM technologies.

Without implemented policies and procedures, the best technology in the world cannot be deemed a trusted system. For example, failing to have **and follow** a policy that specifies the retention schedule may put into question the reliability of your entire system. If, in a litigation situation, a well-meaning manager testifies that although your organization "may" have a retention schedule s/he has his own methodology for when to destroy or retain documents – the credibility of the entire organization to be able to say it has a reliable method for storing and maintaining electronic records may be thrown into question. That means the best technology in the world has not provided your organization with a trusted system. However, if your organization has policies and a documented training program allowing you to demonstrate that manager is an outlier, you should be able to overcome this challenge to the reliability of your ECM system.

Following ANSI/AIIM 25 can help an organization identify all the areas that need to be addressed in order to establish that you have a trusted and reliable system for managing your electronic content.

3. Overview of ECM Technologies

"Many organizations still function almost entirely in a "paper-driven" environment or a combination of both electronic and hardcopy environments. This environment is a direct result of the need of the organization to maintain information on all aspects related to business activities resulting in such a significant increase in retained information that it has become difficult for some organizations to function as effectively as they did prior to the introduction of the internet and wide acceptance of electronic communications. A very important consideration for an organization evaluating or considering EDMS technologies is to first implement the necessary foundational components and then add other functionality as required by the business units, enabling the organization to fully adopt the new technologies and modify the

---

[2] *ARP1 – 2009*, "Introduction."

[3] For more information on this standard or other AIIM or ISO standards, contact Robert Blatt, Chair, AIIM C27 Standardization committee at: blatt@eid-inc.com or Betsy Fanning, Director of AIIM Standards at: bfanning@aiim.org.

business processes as required without adversely impacting day-to-day operations.

To help frame the concepts within this industry best practice document, common terms used to reference these technologies should be discussed. Electronic Document Management Systems (EDMS) have become an all-encompassing term, referring to the integration of various underlying technologies including:

- Document imaging (used to convert hardcopy documents into digital format)
- Document/Library services (used to manage digitally born documents) (Note: Most EDMS systems allows users to use this technology to also manage scanned documents if desired)
- Business Process Management (BPM) / Workflow (used to automate work processes including the creation, routing, track, and management of information being processed)
- Enterprise Report Management (ERM) (used to store electronic formatted reports)
- Forms Processing (used to incorporate interactive forms and manage related forms data)
- Optical Character Recognition (OCR)/Intelligent Character Recognition (ICR) Technologies
- Various applications (also considered add-ons) such as records management applications, legacy system integration tools, etc.

These electronic document management systems provide users with greater access to electronically stored information from common user interfaces typically utilizing industry standard Internet browser technology. One of the primary reasons users prefer this level of technology is the distributed functionality and ability to maintain standard desktop configurations for other office and business related applications.

The structure of EDMS technologies can be viewed as a set of building blocks as noted below in Figure 1 EDMS Technology Building Blocks. The lowest level is the operating system. Database services and Storage Device Drivers are installed onto the server as the second layer. The selection of the database to be used is typically at the discretion of the organization, but has become standardized through the use of Open DataBase Connectivity (ODBC) tools, which has resulted in the database components to be almost considered a "commodity item" rather than a specialized tool.

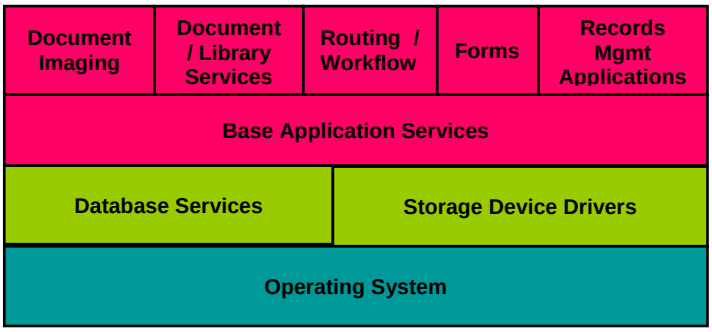

Figure 1 EDMS Technology Building Blocks

The third layer includes the base ECM application components and services provided by the service provider. This layer typically includes the solution configuration tools, application programming interfaces, and application components integrating the core applications components with the database services along with providing the components that integrate the storage environment with the overall solution. The fourth layer incorporates the various core technologies of EDMS technologies Each of these core technologies (with the exception of the operating system layer) are further described in the following sections."[4]

3.1 ECM Technology Providers

As noted in a related technical bulletin in this series, Content/Document Management vendors

[4] *ARP1 – 2009*, sec. 4.1.

within this industry change the names of the core technologies, their pricing strategies and product offerings frequently to keep pace with technology and competition.[5] This bulletin is "vendor-neutral," so the terms used by vendors and product suppliers may be different than industry defined terms; rest assured, however, this bulletin is based upon industry standards such as AIIM ARP 1, and several International Standards, all which are listed in the glossary. The industry best practice, ARP 1, is freely available for download from http://www.aiim.org/standards. Once you understand how the ECM foundation components work together it will be easier to bridge how a specific vendor references their product component with how the industry references that technology.

"Many solution/product vendors that contact you will offer a combination of a systems integrator and value-added reseller (VAR). Over the past many years, these terms have merged together to simply indicate an organization that re-sells and/or integrates the product/solution on behalf of the product supplier.

Many vendors have chosen to be a VAR for some products, then integrate with additional software products to "customize" solutions for their customers. With the advent of open architecture, proprietary systems are seldom seen because they generally do not keep progress with technology advances as well as other systems, are more expensive to maintain, and do not interface well with other systems.

ECM Specialized Consulting Teams can greatly improve the overall project success. As these technologies require internal or specialized resources experienced in change management and how these technologies function, most organizations utilize vendor neutral consultants to become an "internal ECM resource". It is important to note that not all consultants are truly vendor neutral as they may sell products and/or provide services to the software/product suppliers."[6]

4 Foundation Technologies of ECM

4.1 Document Imaging

"Document imaging technologies enable users to scan hardcopy documents into the system and store them in digital format. These technologies enable users to index or enter "metadata" into the system and always utilize some form of storage technology to save the digital version of the document. There are four basic components to document imaging systems:

- input,
- identification (indexing),
- storage, and
- retrieval.

The **input** components typically consist of multiple single-sided (simplex) and/or double-sided (duplex) document scanners (or other input devices such as facsimile). The scanning stations are used to convert hard copy documents into a digital format for subsequent storage and management in the document imaging system. The **identification (indexing)** components allow users to identify (or index) this digital information allowing them to be retrieved at a later date and all types of information required by the end-user organization to fully track all necessary metadata). The **storage** part of the system consists of various components connected to the document management or workflow server and are used to store, retrieve and manage digital information. The **retrieval** part of the system consists of the user issuing a request for information that is then processed by the server. These requests are processed and the information is retrieved from the appropriate storage media connected to the server."[7]

---

[5] Gladwell, Diane. *Managing the Selection of an Electronic Content Management System (ECMS).* Local Government Records Management Technical Bulletin Series. Ed. Robert Blatt. Rancho Cucamonga, CA: MCEF, IIMC, and NAGARA, 2013. Print.

[6] Gladwell.

[7] *ARP1 2009*, 14. More information regarding standards for Document Imaging can be found in *AIIM-ARP1-2009*, section 5.4 et seq.

4.2 Document/Library Services

"Document/Library services technologies enable organizations to manage digitally born documents. Document/Library services applications utilize applets, or thin-clients, to control the authoring, check-in/out, and/or version control of documents being developed, managed or stored. This enables collaborative development when desired along with a mechanism to store/manage digitally born document libraries. The basic capabilities of these technologies include allowing authorized users to:

- Load or import digitally born documents directly into the system,
- Enter relevant metadata associated with the document,
- Create virtual folders linking various documents together,
- Check information in/out of the repository,
- Make changes and check the modified information back into the repository,
- Manage whether original documents are updated or replaced during the update operations,
- Establish security levels for groupings of documents.

The management portion of document/library services technologies include the ability to restrict access to certain documents or groups of documents to only authorized users. Along with security controls, these technologies enable users to be granted different levels of access. For example, the author of a document might only grant read access to all users outside of a specific organization while granting "check-in/out" control to others who are working on updating the document. As the other users prepare to update the document, they would "check" the document out of the library, update the information, and then "check" the document back in.

Document/Library services technologies ensure that any other user attempting to check the document out would 1) not be allowed to check it out, and 2) would be notified that someone already has a copy being updated. Upon completion of the update cycle, the system automatically updates the version number of the document and makes it available to all authorized users.

Similar to Document Imaging, there are four basic processes associated with document/library services technologies (it should be noted that these terms may vary depending on various perspectives such as records management vs. content management vs. archival management, but the underlying functions remain the same with slightly different terminology and/or descriptions):

- import,
- identification (indexing),
- storage, and
- management/retrieval.

The **import** components typically consist of enabling users to import digitally born information into the system. This digitally born information can be any format/structure and can be loaded into the system in original or native format. Digital data does not need to be modified prior to being stored and should include all relevant indexing or metadata associated with the information. The **identification (indexing)** components allow users to identify (or index) this digital information allowing them to be retrieved at a later date and all types of information required by the end-user organization to fully track all necessary metadata). The **storage** part of the system consists of various components connected to the EDMS or workflow server and is used to store, retrieve and manage digital information. The **management/retrieval** part of the system consists of the user issuing a request for information that is then processed by the server. These requests are processed and the information is retrieved from the appropriate storage media connected to the server.

When comparing document imaging and document/library services technologies the following should be taken into account while

determining which technology best meets the needs of the organization:

1. document imaging is the technology used to convert hardcopy documents into digital format,

2. once the document is in digital format, the functionality between these systems is almost identical with the exception of not being able to easily modify the scanned document as can be done with native file formats such as office applications. Etc.

3. document/library services provides a mechanism to easily track document versions while document imaging provides the base capability of simply storing a new document to the system (rather than saving it as a controlled version)."[8]

4.3 Workflow Technologies

"Another module commonly incorporated into the enterprise solution is workflow functionality. Workflow technologies can provide different levels of routing, tracking, and administration. These technologies can be grouped into 3 categories: administrative, ad-hoc, and production. Administrative workflow is typically used by organizations where the processes do not change, or change very infrequently. Ad-hoc workflow provides the ability for a user to create a "work process map" for a specific piece or type of work. Production workflow incorporates administrative workflow and ad-hoc workflow capabilities along with providing extensive tracking and logging capabilities.

When considering production workflow technologies, the organization should consider whether to employ either role-based or user-based workflow technologies. User based technologies require specific users to be assigned to specific tasks, while role-based technologies enable organizations to assign and re-assign users to groups or "roles" which are easily managed. Workflow provides for the automation of business processes and enables users to control the process logic, typically through a graphical user interface (GUI) and other tools. This ability to control the various business processes enables mission-critical, content-centric business applications to operate in an environment otherwise cumbersome to implement and manage. This has resulted in most EDMS vendors offering an integrated workflow engine or integrating the workflow engine, with various workflow products readily available throughout the industry. The difference between these two approaches is whether the product consists of only those components developed by the primary product supplier, or whether the primary product supplier has integrated specialized technologies developed by other suppliers.

In the new approach to organizational networking, workflow is becoming a major tool in the automation of processes and information posting to a web site. In these environments, workflow applications are becoming tightly integrated to legacy applications. The actual integration of workflow and other EDMS technologies has become more prevalent as different coalitions, standards committees, and EDMS vendors have completed development of various standards."[9]

4.4 Record Management Applications

"Another common application organizations implement is referred to as Record Management Applications (RMA). These applications provide configuration and management tools enabling organizations to implement rules associated with identifying the length of time specific types of information (regardless of whether referenced as a document, record, image, etc.) must be kept and the type of disposition that should occur at the end of that retention. There are many other functions and capabilities associated with records management applications, but it is important to recognize that without at least 1 or more of the core underlying EDMS technologies, records management

---

[8] *ARP1 2009,* sec. 4.5.2. More information regarding standards for Document/Library Services can be found in *AIIM-ARP1-2009,* section 5.5 et seq.

[9] *ARP1 2009,* 16. More information regarding standards for Workflow can be found in *AIIM-ARP1-2009,* section 5.6 et seq.

applications will only provide an automated retention schedule.

Records Management Applications (RMAs) are considered to be software used by an organization to manage its records. The RMA's primary management functions are categorizing and locating records and identifying records that are due for disposition. RMA software also locates, retrieves, and disposes of the electronic records that are stored in a repository through integration with relevant core EDMS functions. Any RMA must have at least 1 core EDMS component. Without at least 1 core component the RMA would only be able to manage the policies and not the electronic (or digital) records. It should be noted that RMA functionality is a critical piece of an overall record and/or document management strategy for any organization.

Regardless of whether the organization considers the RMA as a "module" that is integrated into the EDMS foundation or the RMA as the "system" that includes core EDMS functionality, the result is the same: records management policies are implemented to manage electronic (or digital) records contained within at least one of the core EDMS components (i.e, document imaging, document/library services, etc.)"[10]

4.5 ERM (Enterprise Report Management Software)

"Enterprise Report Management (ERM), which was previously known as Computer Output to Laser Disk or COLD, is an integrated software and hardware solution that stores and indexes formatted computer output (pages) on optical disk, magnetic disk, or magnetic tape as an alternative to paper printouts or computer output microfilm (COM). This formatted output consists of point-in-time reports, such as transaction listings of statements and invoices. Once this page output is stored on the ERM subsystem, it can be electronically retrieved, viewed, printed, faxed, and distributed to workstations and host computer terminals within organizations or throughout an enterprise.

While there are many different data types in the computing environment, the type of data which ERM technology is concerned with is typically the result of transactions (data files and database records) being formatted by the application into page-oriented form for printing on paper or computer output microfilm (COM). The structure and format of this output is known. This data is time-period focused—it is a snapshot of an internal system at a given point in time. These reports are often the basis for analysis or comparative reporting and they include the printed record received by users such as a statement or invoice. ERM systems have been designed primarily to handle this formatted output.

Essentially, the ERM process involves two procedures: recording (indexing and storing the data) and retrieving (making the data available to users). Within these two simple procedures, however, lay a myriad of complex tasks. Data must be downloaded or transferred to the ERM system server before it can be processed. The method used to transfer the data from the mainframe/host system to the ERM subsystem will vary depending on the communications capabilities currently in place. Recording consists of writing new documents to the storage media and then making them available for retrieval. Recording speeds vary from system to system and are most critical in high volume systems. The recording process involves:

- transferring the data to the storage subsystem from the host,
- processing the pages from the transferred file (i.e., extracting index keys, compressing, and writing to optical storage), and
- adding the index records to the associated ERM database.

The retrieval process consists of the users accessing the system and selecting the appropriate report, or part of the report, for viewing. The selection of the information to be retrieved is based on information

[10] *ARP1 2009,* sec. 4.5.4. More information regarding standards for RMAs can be found in *AIIM-ARP1-2009,* section 5.7 et seq.

entered, by the user, into the query screen part of the viewing software. After the user selects the report, or part of the report to be viewed, the system retrieves the information, displaying it on the user's workstation."[11]

5. Image Formats

The organization should ensure that all information being scanned, or electronically received is stored in an industry accepted format such as JPEG, JBIG, JPEG 2000, or PDF (and associated variants.

Non-standard or proprietary file formats should not be used. Caution should be exercised if using TIFF. While TIFF is commonly used, there are multiple problems associated with the ability of the application to use non-standard headers, or tags that 1) may not be documented and/or 2) the misuse of other basic headers, or tags. Additionally, TIFF images can be modified without user knowledge though numerous freely available editing tools. Image formats such as PDF-A are non-modifiable through the file format structure along with the use of "checksums" that should be stored in the document management system as an additional method of ensuring that the file has not been altered, modified, or deleted during the information lifecycle.

Non-standard or proprietary formats include any formats used by a single vendor/source and not accepted as a standard file format at either a national or international standards level. Proprietary file formats include but is not limited to:

- File formatting that utilizes "file-wrappers" to encapsulate standard file formats within a nonstandard structure,
- TIFF formats that are not fully documented by the vendor and independently verified by the organization to ensure proprietary information is not contained in any of the headers,
- Non-standard file formats not used by multiple vendors/integrators, etc.

It is important to note that the industry has found that using PDF-A as the output format for any hardcopy conversion to electronic format eliminates many of the commonly seen problems found with TIFF formatted information including: prevents alteration, incorporates the concept of checksums, all information is fully contained, and the PDF-A format is fully standardized and supported by almost every EDMS solution provider, including all the major document imaging solutions currently available (with the exception of smaller solutions that still rely on proprietary methods and concepts)."[12]

6. File Compression

To reduce the overall size of the information being stored and managed, compression technologies are utilized. The most common compression technology used with current ECM products and solutions incorporate "flate" and "de-flate". This compression technology has been shown to be more effective WITHOUT data loss than older technologies.

Other data compression/decompression that can be used include ITU Group 4, LZW, JPEG, JPEG 2000, JBIG, or other output format standards with no proprietary alterations of the algorithms. It is important that the selected compression technology not include extraneous information unsupported by relevant industry standards.

"Users should also be aware that when using proprietary file compression formats, the patent holder may require royalties and/or other fees to be paid on a periodic basis which are usually based on the total number of pages converted into that specific compression format. These licensing/royalty issues do not occur with non-proprietary formats.

There are various compression methodologies available. ISO TS 12033 is a guideline that provides information enabling users to select the appropriate compression technology which the vendor/integrator must support for different types

11 *ARP1 2009,* sec. 4.6.

12 *ARP1 2009,* sec. 5.4.1.4.

of data. The different types of data may include scanned documents, line art, photographs, etc."[13]

7. Planning Activities for Content/Document Management Projects

"You spend months reviewing various technologies, months designing the new system, and many more months developing the solution. Now you have finally finished testing the new system, and you are ready to put it into production. Monday comes along and your users are all eagerly awaiting the new system, which you start up at 8 AM. Promptly at 9 AM the users start reporting small problems, that start building up, and then at 11 AM…… the system crashes, and the users are coming down the hallway "arm-in-arm", chanting your name and asking to speak with you. Have you ever had, or heard of a production rollout occurring in this fashion?

Unfortunately, in our industry the importance of change management activities, detailed roll-out planning and the volume of work associated with document management activities is taken for granted. Document management technologies as defined in the AIIM International Implementation Guidelines…is an all-encompassing term for document imaging, document services, workflow, forms, and other technologies within our industry.[14]

All organizations planning on implementing these technologies should seriously consider the aforementioned planning activities, along with considering the amount of time required to complete them. By way of example, recently a very large California state government agency implemented document imaging and workflow for approximately 250 users, and as described by Sandra Vencill, IWAS Senior Project Manager, Contractors State License Board of the California Department of Consumer Affairs.

'Change management and planning for the internal changes along with the impact of those changes on our users was the most important piece of our system implementation, enabling us to implement our system successfully, with minimal impact to the users. A few examples of the change management activities included in our planning are: user preparation and training, procedural documentation, and ongoing end-user operational and technical support.'

This is a very common perspective of organizations who have implemented production workflow systems with a large number of users. Taking this perspective into consideration, organizations should consider all aspects of rollout planning. Before the organization begins organizing and planning these activities, the following types of questions should be considered:

- Do you want to start-up your system by changing from a manual work process to an automated work process for all users and all work simultaneously?
- Should you remove all existing manual work from the user desktops and insert this into the workflow system prior to Day 1?
- Should you start up the workflow system for portions of the business over a period of days, or weeks?
- Should users begin using the system slowly, completing manual work activities as usual, and getting familiar with the system beginning with a few hours each day, working up to full time, when all the manual work is completed?
- What would be the impact if the system is unable to function for a few hours/days, while "bugs" or "issues" are resolved by the development/implementation team?

There isn't a "right" or "wrong" answer to any of these questions, but rather these questions should be considered, or reviewed, along with other questions and perspectives that will arise during discussions with the users, technical team, and management. The key is to make sure that you

---

[13] *ARP1 2009,* sec. 5.4.2.4.

[14] The AIIM International Implementation Guidelines are available from the AIIM Standards program at no charge at http://standards.aiim.org.

have several, detailed discussions with the user representatives, technical team members, and management during the discussion and planning phases of these activities.

One approach, to rollout planning commonly chosen by end-users and management, is to begin by reviewing the organizational impact. When implementing these technologies, the method and fashion in which users perform their daily, weekly, and monthly work activities will change dramatically. Recognizing that when you implement these technologies, the extent of business process change, and the direct impact on the users and organizational readiness is important. As the organization begins to review the organizational impact and evaluate how the users will accept/adopt the new technologies, issues will quickly be identified.

User acceptance and participation in this level of planning is commonly achieved through having senior management involvement in the project from the beginning, along with the he supervisory and user representatives. These people (as a team) should review requirements, procedures, project plans, and work towards the acceptance of the new technologies and system throughout the entire project.

Getting user buy-in for the project during the design and testing phases is also very important. Experience has shown us that those organizations that utilize "User Champions" (end-users who participate in defining the department needs and share information, project status related information to the other departmental users) have a shorter implementation cycle and faster user/departmental acceptance of these technologies. Also, issues related to training and understanding of the new system, is typically reduced, as many of these items during the initial rollout period are "change management" related, or "training" related issues, and not necessarily technical "problems or issues" in nature.

So, after the organization plans the rollout and prepares the users, the technical team now needs to schedule the actual system implementation. This planning should always include a process of addressing problems/issues that will be encountered, especially during the initial days or weeks after rollout. During this time period, it is very important that the technical team have a thorough plan for correcting and resolving reported problems. Some organizations will try to perform these corrections and updates directly to the production components, while others will work on a development platform, and then only update production after the corrections/updates are verified.

Unless the problem/issue cannot be identified/verified or tested on the development server, it is always considered a very risky proposition to work directly on the production server. A simple update, configuration change, or other update, not properly tested can cause even more problems for the users than the fix/update resolves. Remember that the users will identify and see issues associated with getting their work done. Even if it is a minor issue or problem from the technical perspective, the user frustration can grow quite rapidly, if communication is not good and the user expectations aren't properly managed. Caution should always be encouraged, along with always remembering that user acceptance of the system is always tied directly to the system capability and the ease of implementation along with technology acceptance.

Also, always try to remember that the rollout of these technologies isn't all technical! After the rollout has been completed everyone will enjoy using these technologies, just be patient with the users as this is a new technology for them and a different way of doing business."[15]

8. Recommended Project Phases

AIIM's ARP 1 document, *Recommended Practice: Analysis, Selection, and Implementation of Electronic Document Management Systems (EDMS)*, provides detailed information on various project activities and phases. As this document is

15 http://www.eid-inc.com/white_paper/200210w.pdf

updated at both the national and international levels the AIIM Standards program will provide updates as appropriate.[16]

9. Trusted System and Legal Considerations

"Recognizing that all document management systems manage both electronic documents and records and acknowledging that not all documents become records, organizations may/should (depending on various regulations where appropriate and established) require the same level of system trustworthiness and reliability. Regardless of whether this data is called a "document", "record", or some other term used by the organization, all electronically stored information should be stored in a trusted environment when required and in compliance with the associated record retention schedule/plan.

Taking this into consideration and ensuring that all electronic information is stored and managed in a trustworthy and reliable fashion, compliance with the concepts contained within ISO 15801 and those related to records management policies contained in ISO 15489 Part 1 should be considered. This will ensure that both technical planning, design, and implementation along with records management policies and procedures result in the implementation and operation/management of a trustworthy and reliable document management system for all electronically stored information. It is important to note that a trustworthy system incorporates not only technology but also adherence to documented policies and procedures through all aspects of the design, development, and implementation project phases and be maintainable in an ongoing fashion after rollout into production.

A trusted document management system ensures that all electronically stored information can be considered to be a true and accurate copy of the original information received regardless of the original format. The trusted document management system must ensure that at least two (2) separate copies of the electronically stored information are created meeting, at a minimum, all the following conditions:

(a) The trusted document management system must utilize both hardware and media storage methodologies to prevent unauthorized additions, modifications or deletions during the approved lifecycle of the stored information; and

(b) The trusted document management system must be verifiable through independent audit processes ensuring that there is no plausible way for electronically stored information to be modified, altered, or deleted during the approved information lifecycle; and

(c) The trusted document management system must write at least one copy of the electronic document or record into electronic media that does not permit unauthorized additions, deletions, or changes to the original document and that is to be stored and maintained in a safe and separate location.

It is important to note that trusted document management systems incorporate not only technology, but also require adherence to organizational policies ensuring proper electronic document or record handling, processing as required by the organization (typically documented in the record retention policy and schedule) and electronic document management software or application components….

Expungement

It should be noted that information being expunged needs to follow specific legal rules and does not necessarily require that documents be permanently deleted, but can require that access to documents be permanently removed. Advice from legal counsel should be requested to determine whether permanent removal from accessing documents would meet expungement requirements. ISO

[16] For detailed information on specific project activities, phases and/or updated information on these technologies contact the AIIM Director of Standards, Ms. Betsy Fanning (bfanning@aiim.org) or the Chairperson of the AIIM Implementation Guideline Committee or Chairperson of the ANSI/AIIM C27 Document Management, Mr. Robert Blatt, (blatt@eid-inc.com).

12037, *Electronic imaging – Recommendations for the expungement of information recorded on write-once optical media,*provides specific guidance.

Legal acceptance of records
Evidentiary issues associated with using electronic imaging systems and optical storage technologies need to be considered based upon local legal guidelines. Refer to ISO 12654, *Electronic imaging – Recommendations for the management of electronic recording systems for the recording of documents that may be required as evidence, on WORM optical disk,* for standard guidelines.

Retention requirements
Users and systems designers should consult the organization's established retention requirements set forth in their Records Management Policies and Procedures. The system being implemented should ensure that the system is able to retrieve the information throughout the required document life cycle.

The storage media and its life expectancy rating must be considered, hardware and software obsolescence issues must be evaluated, and a sound migration strategy must be developed to ensure access.

Organizations that do not have current retention requirements should consider developing these documents. These documents enable organizations to manage existing records, and provide a mechanism to automate when documents are to be archived, for how long, what action to take after the retention period is passed, along with numerous other organizational advantages from a management perspective.

Redaction
The process of redaction[17] is elaborate, expensive, and subject to judicial review. It usually involves a careful, word-by-word examination of a document, the identification of the pieces to be "removed," the necessity of showing the location of the removed pieces, the inability of the document viewer to discover the redacted content, and the supervisory review and approval of the redaction - all with the recordkeeping to prove that the redaction was appropriate and conducted according to proper procedure.

Therefore, the redaction process is usually done in a highly-controlled local setting. Redaction process software could be external to the document management system."[18]

9.1 Security and Location of the Electronic Stored Information (ESI)

In addition to prohibiting access from outside the organization, consideration must have been given to who has access to the ESI from within the organization and where and how it is stored.

To meet the best practices, as identified in section 5.3.1 of *ARP-1 (2009),* a trusted system must generate two separate copies of ESI when it is created. And, that "[t]he trusted document management system must write at least one copy of the electronic document or record into electronic media that does not permit additions, deletions, or changes to the original document and that is to be stored and maintained in a safe and separate location."[19]

For example, it would not be within industry best practices to have a system that, when a hard copy document is scanned, generates only one copy that is later backed up through the disaster recovery protocol. Industry experts believe that by

[17] "Redaction" refers to a process by which parts of a document are kept from disclosure. Documents might contain pieces of information that are protected by law from being revealed, e.g., because they contain privacy identifiers or trade secrets or other privileged information. The parts might be such snippets as the name of a person, a Social Security Number, or entire paragraphs that reveal trade secrets. In many redactions the rendition of the redacted document, whether hard or soft copy, will show a black bar through the space where the redacted content was located.

[18] *ARP1 2009,* sec. 5.3.3 – 5.3.3.4. Additional information related to all trustworthiness and reliability are documented in ISO 15801.

[19] *ARP1 2009,* sed. 5.3.3.

generating multiple copies to separate locations at the time the document is created substantially reduces the risk of failure, loss, or alteration of the original document so as to better protect the integrity of the information stored.

9.2 Integrity of the Information

By requiring the storage of information in a system that uses hardware and media methodologies to prevent "additions, modifications, or deletion" *ARP-1 (2009)* has captured the concept that users must have confidence in the accuracy of the stored information. A document management system that does not provide the users with confidence that the information put into the system may be regenerated accurately is of diminished or no value. In fact, it may end up with a negative value if it requires additional costs to provide that needed assurance to accuracy. For example, in the context of producing records for a litigation request, someone in the entity must be able to verify that the documents generated were "true and accurate" copies of the documents maintained. Absent the level of confidence generated by following industry best practices, companies will have an uphill battle demonstrating their documents are accurate, unless they also maintained the original documents from which the ESI copy was produced is compared. This would result in inefficiencies in the business including increased costs in storage and time to review and compare the documents.

The standard-setting committee that developed *ARP-1 (2009)* left latitude to designers to create a system utilizing hardware, software and media that best meets the needs the particular business organization. For example, an entity that modifies its original document throughout the course of the document lifecycle by appending information may have an entirely different design to its system than an entity that simply needs to maintain the original information. Regardless of how each system is crafted, it must consider preserving the integrity from the original source.[20]

[20] AIIM Implementation Guidelines Committee can be found at http://www.aiim.org/standards or by contacting the AIIM committee chair and ISO Project leader at blatt@eid-inc.com .

9.3 Audit Trails & Historical Data

Another important aspect to the overall document management system is having an independently verifiable audit trail that can demonstrate the ESI has not been altered inappropriately. Once again, this provides added confidence to the end user that the documents maintained and/or generated from the system are accurate. The level of information that is available within document management systems can be at the level of who opened and printed, to who took what actions including workflow notations and/or routing to other users for review/processing. Being able to demonstrate reliability is directly related to both security AND appropriate levels of all forms of historical data that comprise the various audit trails.

9.4 Policies and Procedures

The final component identified as a best practice in a trusted document management system is the creation of and use of policies and procedures to support the system. It does little good for a document management system to be created if there are no guidelines for the type of information to be stored or identifying what methodology will be used to generate ESI.

For example, a business that transacts with clients via email should consider a policy for when an email should be stored in the document management system, giving guidelines for how employees should know whether it should be stored and how to implement the policies. Absent those types of policies, employees will have little guidance and differing practices for when email communications are stored in the system will develop throughout the organization.

10. Characteristics of Trusted Storage

"For a storage medium to be trusted as a component in an ECM system, it must have several key characteristics. It is very important to note that storing on media that is not directly controlled by the content/document management solution almost always fails the test of Trusted Solutions. Another

key item is that the storage media MUST prevent any unauthorized access, modification and/or deletion during the information life-cycle.

This section describes the key requirements for trusted storage sub-systems.

10.1 Content is Accurately and Permanently Stored

For storage to support ECM, it must be trusted to store exactly what is sent to it from ECM. Any alteration of content may make the data inaccessible, unusable by the application that generated it, or make the content unacceptable or inadmissible as a copy of record.

The original import of the content is an important risk vector. If the file is not properly stored, and then the source content is deleted, it could be permanently lost. Storage mediums generally do not have issues with storing exactly the bits sent to it. This is more likely to be encountered as a challenge if a storage solution rewrites the file with extra metadata, so this type of implementation can be problematic.

Procedures to verify the correct capture of content must exist. They might involve checking error logs, or the sampling of content in the ECM system to make sure capture is occurring correctly.

The reliability of the storage medium is essential. Replication of content to other mediums can help address the business risks of failure. The storage medium may have the only copy of content until that replication occurs, and uncertainty of the content's authenticity is introduced if anything happens to the data on its original storage location, as these changes might then be replicated.

10.2 Content Cannot Be Erased, Rewritten, or Moved

Once the content arrives at the storage medium, the storage medium is responsible for guaranteeing the file continues to exist. It is also responsible for preventing unauthorized unlogged alteration or deletion of files. Another issue is the moving of content without the ECM System being aware of the move. This typically makes content irretrievable by the software, and potentially breaks a custody chain.

Unauthorized content deletion is a significant problem, but undetectable alteration of content is worse. If it is possible to alter content undetectably, the authenticity of all records is made questionable. The risk of content on disk being deleted or altered unintentionally is best addressed by procedures to identify and protect data that will be kept as a record. These procedures should include the application and review of access security controls, and the verification of data before it is out of scope of available backups.

Some implementations store a retention period at the storage medium, preventing deletion until a point in time. This can be helpful, but doesn't replace responsible management of retention.

10.3 Access to Content is Secured and Auditable

Access to content must be authenticated and audited. Many of these responsibilities are owned by the ECM system, the network authentication store, and the security policies and configuration. The storage medium still has the responsibility for permitting access only from authenticated users.

10.4 Storage Provides Accurate Status and Error Information

Software that writes content to a storage medium should check for acknowledgment from storage that the content was successfully written to the storage medium, or wasn't. The storage medium must provide accurate and complete status to the interface it communicates with, typically the operating system handling the I/O operation. This permits the software storing content to take appropriate actions, such as logging the failure, prompting the user, or attempting to copy the data to another location.

When status information is not accurately communicated, the ECM software may record the

successful (and incorrect) writing of content to storage. This can result in lost data, and makes the authenticity of the other records in the system questionable.

10.5 File Metadata is Maintained

Part of demonstrating authenticity is to verify the age of the content. File-level metadata from the storage system is a way to verify the information in the ECM system. Without file-level metadata to support the claims of the ECM system, it is difficult to prove that the content on the storage medium is the same content originally stored by the ECM system.

10.6 Audit Trails and Verification

To prove authenticity, to verify that security controls and processes are working correctly, and to show a custody chain, an audit trail of access to content is necessary. These audit trails at a minimum should log creation dates, changes to metadata, any change in physical location, and any re-indexing that has occurred."[21]

11. Change Management & Communication Best Practices

The key to success in implementing EDMS lies in integrating the key elements described so far: Technology, Readiness, Operations, and Culture. The best way to accomplish this integration is through the active participation and involvement of users and management.

Champion User Participation

"Participation from all levels of employees in the implementation process is an underlying theme of the following recommendations. At a minimum, key users (also referred to as "champion users") should be identified throughout the organization. These "champion users" are typically senior or lead users who can provide input and feedback via a bi-directional communication model enabling the EDMS team and the users throughout the organization to be involved in all appropriate aspects of the analysis, design, and implementation project phases.

The benefits of employee participation are increased motivation, higher productivity, and improved quality. In one study in which resistance to work changes was lower in groups that participated in making those changes, researchers identified two key points:

a) Participation is a necessary but not sufficient means of reducing resistance.
b) Participation is a feeling of involvement on the part of people, not just the mechanical act of taking part in discussions.

Organizations that have left the 'champion users' out of the planning, problem solving, analysis, and redesign or that have only marginally involved employees through random conversations and presentations have been unable to tie together the four key elements: Technology, Readiness, Operations, and Culture. One of the best ways to ensure participation is through a Champion User Team.

Selected by senior management, this team should comprise representative individuals from all levels and all key job functions as well as members of the technical project staff. The goal of this group is to jointly design the new work processes and jobs to best utilize the EDMS and human resources. The formation of this team will alleviate many of the us-versus-them problems that arise when the technical staff, records management team, or end-user representatives work separately from the other portions of the business unit in designing a system.

This Champion User Team should begin its work with a one-day or two-day training session that would review the following topics:

- team charter, roles, and responsibilities;
- project objectives and goals;
- change parameters established by senior management;

[21] AIIM C21 – Advanced Data storage standards committee. For more information contact the AIIM C21 committee chair and ISO TC/171 SC1/WG8 Convenor, Robert Blatt at blatt@eid-inc.com or Betsy Fanning, Director of Standards at bfanning@aiim.org .

- methodology for work redesign that looks at both the social and technical aspects of work;
- problem-solving techniques such as brainstorming, flow charting, cause-and-effect diagraming, and the like;
- development of effective teams;
- effective team behaviours and,
- use of new technologies as support for the project.

The Champion User Team should continue to meet on a regular basis to accomplish the following tasks:

a) **User analysis:** identify users, determine the extent to which their needs are being met, and identify actions that can be taken to increase user satisfaction.
b) **Technical analysis:** document how work flows, where problems occur, and where these problems are first discovered. Also determine how imaging can impact the current workflows and which variances will be eliminated or reduced by imaging.
c) **Work redesign:** according to information already gathered, rethink new workflow designs.
d) **System design:** according to information already gathered, finalize system specifications.
e) **Organization design:** according to new workflows and system capabilities, determine structural changes that need to be made in the organization, if any.
f) **Implementation plan:** identify the steps and resources required to move from the current organization to the new image-enabled organization.

Change Management Program

Once the data have been gathered and analyzed, a targeted change management program can be created. It is the follow-through on carrying out the change management program that will be most critical to the success of the implementation process. The change management program should look for additional opportunities to involve key users or "champion users" in all aspects of the EDMS project. This can be done by establishing User Involvement Teams. The teams would be created to develop strategies for managing the key issues that were identified through the organizational assessment. For example, separate teams might look at issues such as communications, training, and policies and procedures. Since many of the team members may not have had opportunities to help shape their organizations, the teams such as the Design Team will need some initial training in team building, problem solving, and meeting management.

While each organization will develop a change management program that is unique to its own situation, three broad areas that every organization will need to address are communication, training, and job design.

Communication

In most organizations, communication is not as good as it should be. While this is not usually detrimental to the business, lack of communication at critical times of change can lead to failure of the change effort.

The following are suggestions about communicating during times of transition:

a) Fully describe the problem previously identified by end-users and management.
b) Acknowledge input from users that were interviewed and participated in the process baselining activity.
c) Identify the following:
    1) What information is difficult to find?
    2) What challenges do the users encounter with the current process?
    3) Who will benefit from technology based change?
    4) What types of non-technology based change is desired or required?
    5) What will actually change?
    6) What secondary changes will occur?
    7) What will be changing for the organization as a whole?
d) Compensate for losses. For example, in EDMS environments, many employees feel a loss of socialization with their peers or feel that the use of these technologies will increase management oversight and force

users to "account for their time". While this may be an objective, how this is presented to the end-users will greatly affect their interest and willingness to participate in any type of change.

e) Communicate as much as possible. It is always better to over communicate the project goals, objectives, and status, rather than under communicate the plan and what is scheduled to occur at each point of the project.

f) Emphasize and demonstrate the active and direct support of all levels of the management team.

Some successful EDMS installations have relied on monthly or quarterly meetings, newsletters, electronic mail updates, WEB site communication pages, hot lines, and suggestion boxes. People in each organization will need to determine what will work in their environment.

Senior Management Participation

The most important contribution senior executives can make early in the project life cycle is to participate in a strategic planning session. The purpose of this session is to clearly articulate desired project goals and objectives and the desired organizational change. Most organizations develop a Technology Strategy which includes EDMS. Usually this is tied to a Business Strategy. The third component of this strategic triangle is the Organizational Change Strategy. In many cases, this third critical strategy is non-existent. Failure to articulate an Organizational Change Strategy can lead to failure to manage the human and organizational impact of EDMS. Trying to manage the human and organizational issues without a coherent strategy will result in an unfocused and ineffective change management effort.

Strategic Planning Session should focus on answering the following questions:

- What aspects of our business culture are effective?
- What aspects of our business culture are ineffective with regard to EDMS implementation?
- In what ways will EDMS impact our employees and key external organizations?
- What structural changes in the organization are likely to result from EDMS?
- How much change do we want in this organization? When examining the continuum of control versus commitment, where are we now with regard to management style? Where do we want to be?
- What technology based change is appropriate for the organization?
- How should the project be phased to allow adequate time for change management and organizational acceptance of the selected technologies?
- Do we simply want to automate existing processes or do we want to fundamentally change workflows?

The outcome of these meetings should be clearly stated objectives for change management and a set of guidelines for change that can be used by implementation teams. For example, are there any processes or procedures that cannot be changed due to regulatory requirements? Can we eliminate certain positions; alternatively, must all job titles remain in the organization? Can we redirect resources? Do we want to increase employee participation a little bit? Alternatively, do we want to move toward self-managed work teams?"[26]

12. Document Classification & Indexing Model (Taxonomies)

"An important aspect of any ECM effort is to establish an enterprise wide document classification and indexing methodology. This effort should include meeting with all appropriate staff identified by the organization as being representative of how the ECM solution would be used to discuss existing document classifications and to prepare a full classification plan that can be implemented by the selected solution integrator/software provider. It is important to note that when implementing an enterprise foundation the resultant document classification and indexing methodology should:

---

[26] *ARP 1 2009*, sec. 6.1.1 – 6.1.4.

- be flexible and able to be updated as required by the organization without adversely affecting
- documents already stored by the solution;
- fully addresses records retention issues and schedules;
- enable staff to easily locate and retrieve documents;
- support the ability to implement document security;
- provide logical information groupings; and
- support the use of 'virtual folders' where staff can search and organize documents during the
- retrieval process without forcing document replication or copying.

This classification should be fully compliant with the organization records retention schedules and link to the records management classification and structure. This will enable the business side of the organization to further manage the business documents and records as required by the records management team.

Business objectives and requirements:
The business objectives, functional requirements and expectations should be clearly defined. The business / functional requirement documentation should include both technology driven and nontechnology driven requirements and detailed information related to:

- Business objectives of the project;
- Business functional requirements; and
- Business expectations.

This document should contain specific information related to current and near-future business needs and requirements identified through the business baseline activities, and interviews with business, technical, and management team members of the organization. Additionally, the definition of critical success factors (CSF) associated with how the resultant solution would be evaluated to determine overall achievement of these objectives and requirements should be clearly defined. Critical success factors are those items considered to be either a business or technical requirement for the organization. The CSF's should enable the organization and the vendor to identify those areas of critical importance related to the successful implementation of the desired technologies."[27]

13. Common Critical Success Factors

"Common examples of critical success factors from both a business and technical perspective include:

Business related goals:

- Improved service: Users need the ability to quickly access and review information managed by the document imaging and workflow system.
- Ability to track and monitor work activities: The system should enable the users to track all ongoing work including the ability to re-assign work from one user to another. This tracking capability will enable the organization to implement workload leveling when appropriate.
- Centralized historical information between organizations: The system should enable the organization to maintain centralized history related to all activities associated with the client/constituent. This history centralization should include both system generated activities (i.e., date scanned, date routed, etc.) and user generated information such as notes taken during telephone conversations. The users should only have access to information allowed by their security access. The system should restrict access to information required by higher levels of security.
- Increased efficiency of available resources: The organization should be able to use the selected technologies to support ongoing business activities. The selected technology should enable users to decrease time spent on paper and file handling activities including stamping, stapling, copying, delivering, and filing documents, and increase time in the areas of work processing.

[27] *ARP1 2009*, sec. 6.6 – 6.7.

- Satisfy organizational and/or government regulations pertaining to document retention: The use of electronic data storage must adhere to any laws and/or regulations covering the storage, retention, and retrieval of information on electronic storage media.
- Decreased storage costs: The solution must provide the ability to use optical storage technology to reduce the overall cost of storage and retrieval of all "hardcopy" information.
- Decreased costs for manual document management: The cost for manual document management should be reduced along with an increase in the ability to provide improved service at a lower cost per request.
- Simplified user access to application, work-order, and other data: The overall solution must enable the users to quickly select and access the desired information without using highly complex user interfaces or tools. The user interface needs to be easy-to-use by the various system users.

Technical goals:

- Scalability: The system must be fully scalable, allowing for an increase of the number of users and volumes of data without replacing primary system components. This scalability must be in the areas of increased memory, disk storage, optical storage, CPU speed and size, etc.
- Migration path: A clearly defined migration path must be fully supported by the proposed solution.
- This migration path must provide for the integration of new document management technologies to ensure proper integration without adversely affecting the proposed solution and/or data managed by the existing system(s).
- Modularity: The various client-based applications must be modular allowing for implementation of additional functionality without adversely affecting the overall system solution. This includes the ability to add routing; "virtual" file folders, high-volume printing, automated fax services, workload distribution, monitoring, etc.
- Browser based access: The system must fully support browser based technology where the various web servers will provide all the necessary mechanisms to store and retrieve information requested by the user, system level security for both users and data, and associated system management functions. All applications must be fully integrated to prevent redundant hardware and software on both the workstation and web server platforms.
- Use industry standard components (no proprietary architectures allowed): The associated components within the solution must be commonly available throughout the content / document management industries, be fully supported by the selected product supplier, and have full user and/or development documentation and libraries."[28]

14. Final Note

To get more detailed information on ECM related standards and best practices, the AIIM International Standardization program is a great source of information. For more information on the various content/document management related standards and best practices, contact Robert Blatt, Chair Implementation Guidelines at IG@eid-inc.com or Betsy Fanning, Director of Standards at bfanning@aiim.org.

[28] *ARP1 2009*, sec. 6.7.

15. Glossary

In Laymen's Terminology from AIIM TR2 - Vocabulary[29]

ANSI: An organization that oversees the development, promulgation and use of industry standards (American National Standards Institute) (http://www.ansi.org).

API: Application Programming Interface - software that allows different software components to communicate with each other. Also see Developer's Toolkit.

ASCII Text: Letters, words, intelligible characters. ASCII (American Standard Code for Information Interchange) is an adopted standard code to represent a character (e.g. A = 00010001 in ASCII). **Text** - what word processors use.

ASP: Application Service Provider. A private business that provides computer-based services to customers over a network. (think "Rent") Also see Cloud Computing, Hosted Solutions, and SaaS.

Backfile (conversion strategy): The process of scanning, indexing and quality checking of documents, most commonly paper-based but including microfilm and microfiche, then storing them in a digital format. Also see Day-forward.

Boolean Logic: And/or/not searches (i.e. a full text search for "Diane and raise" will result in a smaller number of hits than "Diane or Raise").

Browser: A web browser is a software application for retrieving, presenting, and traversing information resources on the World Wide Web.

BPM: Business Process Mapping (as used in this bulletin.) Developing a graphic "map" or illustration of activities involved in defining exactly what is done, who is responsible, etc.

Cloud Computing: The "Cloud" is used as a metaphor for the Internet, based on how the Internet is depicted in computer network diagrams. Private companies using the "cloud" can provide storage capacity, software and services over the Internet. Think "rent" – also see ASP, Hosted Solutions, and SaaS.

CD: Compact Disk. Optical media that uses lasers to create pits on the media. **CD-r** is unalterable media. Holds up to 4 filing cabinets scanned of records; comparatively slow access speed. .

Client/Server ("Thick Client"): Network configurations that rely on a "server" computer that provides programs and/or files to its "client" computers (users) as needed.

COLD: Computer Output to Laser Disk (now called ERM). File uploads, usually in ASCII text, then overlaid with your form template (think financial reports – the data is columnized and paginated.)

Concurrent (Software) Licensing: Licenses are "checked out" by various employees up to the limit of the licenses purchased. Also see Named User (Software) Licensing.

Data: Information.

Data Mining: Extraction of detailed and summary data / records.

Database: An organized collection of data (information).

Day-forward (Conversion Strategy): Place only newly created or received records in the system (also see *Backfile*).

Deskewing: Straightening out a page that was scanned in crooked.

Despeckling: Removing noise (speckles of dirt) from an image.

---

[29] © 2012. Reprinted with permission.

Developer's Toolkit: software that allows different software components to communicate with each other, and can provide customization of a software product. Also see API.

DOD 5015.2 – The security standard for computer systems that is used by The United States Department of Defense, and is used as a benchmark by many systems.

DPI: Dots Per Inch is the number of individual dots that can be placed in a line within one inch (e.g., 300 dpi is 300 dots per inch.) The higher the DPI number (resolution,) the better the quality of the image, and the larger file size it will be.

DVD: Digital Video Disk. Optical media that uses lasers to create pits on the media. Data on a DVD-R cannot be changed, whereas a DVD-RW (rewritable DVD) can be rewritten multiple times. Holds up to 64 filing cabinets of scanned records; 7 times as fast as CD.

ECMS: Enterprise Content Management System. Tools used to capture, manage, store, preserve, and deliver records ("content").

EDMS: Enterprise Document Management System. Tools used to capture, manage, store, preserve, and deliver records ("documents")

ERM: Enterprise Report Management (was called "COLD"). Used to store electronic formatted reports (for example, from your financial system) so it has the column headings, page breaks and other formats to make the report easy to read.

ESI: Electronic Stored Information (data / records)

Expungement: Process by which record is destroyed, deleted, or sealed from the repository.

Form Drop Out: Usually used in conjunction with ICR; to remove the form prior to recognizing handwriting.

Full-Text Search (Engine): Searches through every word in a database (an engine is another name for a program).

Fuzzy Searches: The ability to find words with similar spellings (rather than the exact word match.)

Hardware: Physical equipment (computer, printer, etc.)

Hierarchial Folder Structure: Graphically representing files or groups of files so they look like folders in the Windows operating system.

Hosted: A way of providing programs over intranets / internets via an outsourced e-commerce provider. (Think rent.) Also see ASP, Cloud Computing, and SaaS.

ICR: Intelligent Character Recognition. Turns a handwritten image into ASCII text, which can be used by other software programs.

ILM: Information Lifecycle Management. Practice of applying certain policies to the effective management of information throughout its useful life (think retention).

Image/Raster Image/Bit-Mapped Image: A graphic representation of a page by millions of dots (pixels).

Internet: A global network of interconnected computers, enabling users to share information. (think "Cloud Computing")

Intranet / "Thin Client": Enhanced LAN or WAN installed within an organization; provides functions similar to the Internet. Has "firewalls" to protect against Internet hackers. Internet/Intranet ECMS configurations are called "thin clients," and although they are easier to deploy and support, "thin clients" may not have all the functionality of a "thick client."

ISO: An organization that develops and/or approves industry standards (International Standards Organization).

JPEG: Joint Photographic Experts Group. Another common **image** (picture) format that is compressed - has "flavors" or variations.

LAN: Local Area Network (computers linked together) usually in the same building.

Magnetic Media: Hard drives, fileservers, USB keys, diskettes, cassette tapes, RAID systems. All use selective magnetization of portions of the surface to record data, and can be re-written.

Media: What a record is stored on: Paper, magnetic (computers, tapes, USB keys), microforms, optical disks (WORM / CD / DVD), etc.

Metadata: Index information added to describe the document (e.g. date, document type, document number). Also called template data, field data, index data, etc.

MFD's: Multi-Function Devices. Typically a combination of copier / scanner / printer / fax services.

Named User Licensing: Software licensed to each user (each user has their own license, versus concurrent licensing, where there are licenses that any one user can use up to the maximum allowed.)

NAS: Network-Attached Storage is computer data storage connected to a computer network; utilizes resources more efficiently than traditional file servers (thereby increasing capacity and speed.) Also see SAN.

Network: Two or more computers linked together.

OCR: Optical Character Recognition. Turns a typed image into ASCII text, which can be used by other (e.g. word processing) programs. (Zone OCR targets a selected area on a document and OCRs only that area.)

ODBC or OLE: Open Data Base Connectivity (ODBC); Object Linking and Embedding (OLE) Standard database formats that allows information to be written in one program, then copied into another program and/or allows embedding and linking of documents.

Operating System (OS): The base program that allows other programs to run (e.g., Windows 7).

Optical Disk: CD, DVD, or WORM; uses lasers to create pits in the strata of the media.

PDF: Portable Document Format. May store some elements outside the file (e.g., fonts, hyperlinks, some graphics.) Developed by Adobe for their Acrobat product.

PDF-A (or PDF/A): An archival standard that requires all elements to be stored inside the file, which ensures that documents can be reproduced the exact same way in years to come (ISO 19005).

Portal: A framework to provide a single point of access to a variety of information and tools (Think of your Intranet page used by employees)

Peripheral: Anything connected to your computer (Printers, keyboards, etc.)

Program/Software/Application: A series of instructions. This tells the computer what to do.

Protocol: Format to communicate between two computers. (e.g. TCP/IP).

RAID: Random (or Redundant) Array of Independent (or Inexpensive) Disks. Used to increase the capacity of traditional file servers. By combining several large hard drives, you can configure a file server with greater capacity that is protected against failures. There are different "levels" of RAID that have different capabilities.

Re-Writable Optical Disk/Magneto-Optical Media/Phase-Change Optical Media: A type of optical disk that allows you to change the data and re-write.

RMA (Records Management Applications) Provides management tools for electronic and hardcopy records (e.g. boxes of paper records stored at an inactive records center.) Can combine ECMS / document imaging capabilities.

SaaS: Software as a Service. A way of providing programs over Intranets / Internets via an outsourced e-commerce provider. (Think rent.) Also see ASP, Cloud Computing, and Hosted Solutions.

SAN: Storage Area Network is a dedicated network that provides access to consolidated, block level data storage, and utilizes resources more efficiently than fileservers (thereby increasing capacity and speed.) Also see NAS.

SharePoint: Collaboration Software from Microsoft. SharePoint is focused on shared workspaces; users can manipulate or interact with lists and document libraries (sites.) SharePoint can be used to provide intranet portals, document & file management, collaboration, social networks, extranets, websites, enterprise search, and business intelligence. It also has capabilities around system integration, process integration, and workflow automation.

Social Software: Web-based software programs that allow users to interact and share data with other users. Facebook, and Twitter are just two examples.

Storage / Archive and Retrieval (Document Imaging): A type of ECMS technology that stores, then retrieves documents. In many cases, scanning converts the hardcopy document into a digital image (via scanning,) and metadata (indexing) is attached to the image so it can be searched and retrieved

Taxonomy: Classifying Information / Naming Conventions (controlled vocabulary.) Encompasses description, identification, nomenclature, and classification

Text / ASCII Text: Letters, words, intelligible characters. ASCII is an adopted standard code to represent a character (e.g. A = 00010001 in ASCII).

Text Mining: Locating text that contains specific terms in an electronic record (Full Text Searches.)

Thin Client: Software deployed over an "Intranet" and accessed by launching your browser (via your organizations Intranet.)

Thick Client: Software deployed over your internal network in a client / server configuration; software must be installed at the client workstation.

TIFF: Tagged Image File Format: A file format for storing images, popular for ECMS repositories. As of 2009, it is under the control of Adobe - has "flavors" or variations.

Unstructured Data: Data files stored on computers (in the Operating System, not in a database). Word documents, Excel Documents, e-mail, voice mail, IM, etc. not stored in a repository.

URL: Universal Resource Locator (your e-mail address)

USB Flash Drives / Keys / Thumb Drives: A data storage device that includes flash memory with an integrated Universal Serial Bus (USB) interface that uses magnetic media (re-writable.)

WAN: Wide Area Network (computers linked together) located in different buildings.

WCM: Web Content Management - Used for creating and managing HTML content / Internet Web Sites.

Workflow: Routing and scheduling documents. A rudimentary workflow system can be achieved with e-mail; a true workflow system has functions such as process and job definition, scheduling, reminders, auditing, parallel and ad hoc routing, event monitoring and simulation.

WORM: Write Once, Read Many Times. Archival media (images cannot be altered).

WORM Optical Disk: Uses Optical Disk (unalterable) media.

WORM Tape: These systems use a combination of software and hardware techniques to make the tapes both non-rewritable and non-erasable.

16. Resources

AIIM (Association for Information and Image Management,) http://www.aiim.org

AIIM ARP1-2009 – *Analysis, Selection & Implementation of Electronic Document Management Systems (EDMS)*

American Association for State & Local History (AASLH), http://www.aaslh.org

Association for Information and Image Management (AIIM) http://www.aiim.org

Association of Records Managers and Administrator, International (ARMA, International,) http://www.arma.org

Council of State Archivists (CoSA), www.statearchivists.org

Dearstyne, Bruce ed. *Leading and Managing Archives and Records Programs: Strategy for Success.* New York: Neal-Schuman, 2008. Print

Freeman, Robert, Mark R. W. LeMahieu, Jason Sanders, David O. Stephens, Julian L. Mims III (editor), *Electronic Records Management.* International City-County Management Association (ICMA), 2006. Print

International Institute of Municipal Clerks (IIMC) http://www.iimc.com

International Organization for Standardization (ISO), http://www.iso.org/iso

ISO/ 10244, Document management – Business process/workflow baselining and analysis

ISO/TS 12032, *Document imaging – Statistical sampling for document images*

ISO/TR 12033, *Electronic imaging – Guidance of document image compression methods*

ISO/TR 12037:1998, *Electronic imaging – Recommendations for the expungement of information recorded on write-once optical media*

ISO 12651:1999, *Electronic imaging – Vocabulary*

ISO 12651-1, *Electronic document management – Vocabulary – Part 1: Electronic document imaging*

ISO 12653-1:2000, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 1: Characteristics*

ISO 12653-2:2000, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 2: Method of use*

ISO/TR 12654:1997, *Electronic imaging – Recommendations for the management of electronic recording systems for recording of documents that may be required as evidence, on WORM optical disk*

ISO/TR 14105:2001, *Electronic imaging – Human and organizational issues for successful electronic image management (EIM) implementation*

ISO/TR 15801, *Electronic imaging – Information stored electronically – legal admissibility and evidential weight*

ISO 15489-1, *Information and Documentation-Records Management-Part 1: General*

ISO/TR 15489-2, *Information and Documentation-Records Management-Part 2: Guidelines*

ISO 17799, *Information technology -- Security techniques -- Code of practice for information security management.*

National Archives and Records Administration (NARA), http://www.archives.gov

National Association of Government Archives (NAGARA), http://www.nagara,org

National Historical Publications and Records Commission (NHPRC), http://www.archives.gov/nhprc