TECHNICAL REPORT

ISO/TR 22957

First edition
2009-12-15

Document management — Analysis, selection and implementation of electronic document management systems (EDMS)

Gestion de documents — Analyse, choix et mise en œuvre de systèmes de gestion de documents électroniques (EDMS)

Reference number
ISO/TR 22957:2009(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword iv

1 Scope 1

2 Normative references 1

3 Terms, definitions and abbreviated terms 1
3.1 Terms and definitions 1
3.2 Abbreviated terms 1

4 Electronic document management technologies 3
4.1 General 3
4.2 Database services 4
4.3 EDMS application services 4
4.4 Core technologies and application specific modules 4
4.5 Document imaging technologies 4
4.6 Document or library services technologies 5
4.7 Workflow technologies 6
4.8 ERM technologies 7
4.9 Forms processing 8
4.10 Optical and intelligent character recognition 8
4.11 Records management applications 8

5 EDMS guidelines and standards 9
5.1 General considerations 9
5.2 Trusted EDMS technologies 9
5.3 Industry guidelines 10
5.4 Document imaging 14
5.5 Document/library services 18
5.6 Workflow 19
5.7 COLD/ERM 23

6 Best practices associated with EDMS project phases/activities 23
6.1 General 23
6.2 Process/procedure baselining 23
6.3 Anticipated processes/procedures 26
6.4 Technology requirements definition 26
6.5 Document classification and indexing model 26
6.6 Business objectives and requirements 27
6.7 Technology evaluation guidelines 29
6.8 Forms review and design considerations 30
6.9 Legacy data/document conversion methodology considerations 31
6.10 Procurement document preparation 32
6.11 Solution/product evaluation guidelines 32
6.12 Project planning and execution 34
6.13 System, unit testing, and project monitoring 34
6.14 Acceptance testing criteria 35
6.15 Rollout planning 36
6.16 Business practices documentation 36
6.17 EDMS project activities best practices 36

Annex A (informative) Guidelines and standards 37

Annex B (informative) Summary listing of industry best practices associated with EDMS Project phases/activities 40

Bibliography 41

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 22957 was prepared by Technical Committee ISO/TC 171, *Document management applications*, Subcommittee SC 2, *Application issues*.

Introduction

This Technical Report provides detailed information associated with the analysis, selection and implementation procedures associated with electronic document management systems (EDMS). The development of this Technical Report is a result of organizational requests to receive vendor-neutral industry information associated with technology standards, technical reports, guidelines and best practices related to project activities.

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 used, incorporating additional levels of functionality, and they are very rarely a result of an entirely new core technology. This is important to note, as the core EDMS technologies are constantly maturing and solution providers are not only identifying new approaches to addressing organizational issues and requirements, but also expanding the use of these technologies into areas previously unconsidered.

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 hard copy document scanning into digital images in the early 1980s, to the management of any digital or electronic document that today is referred to as electronically stored information.

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

The first section of this Technical Report provides detailed information describing each of these technologies, and how they operate and inter-operate.

The second section of this Technical Report provides detailed information associated with currently available industry standards and technical reports.

The third section of this Technical Report provides detailed information related to industry best practices associated with all the customary project phases for EDMS technology analysis, selection and implementation. These project activities are considered to be industry best practices. It has been demonstrated over the past 10 years that organizations following all the recommended steps and activities have a much greater level of project success while greatly decreasing, and in most cases, eliminating unnecessary technologies, user licenses, etc. This is very important, especially with most organizations carefully examining all expenditures related to technology procurements.

This Technical Report provides detailed guidance to organizations considering the use of any of those technologies that comprise EDMS (document imaging, document/library services, routing//workflow, records management applications, forms management, enterprise report management, etc.). A complete records management programme is critical to any organization and is integral to any complete and thorough management plan associated with electronic information, regardless of whether it is internally referred to as a “document”, “record”, “audio”, “video”, etc., by the organization.

All relevant project steps, tasks and activities contained within this Technical Report, together with compliance with relevant industry standards and guidelines, should be examined and “certified” by the technical implementation team as being in compliance with these industry best practices, thereby ensuring, especially for organizations that are required by government codes and/or regulations, that industry best practices, guidelines, and/or standards established by ANSI, AIIM, and/or ISO are followed.

The term “electronic document management” used throughout this Technical Report is intended as an “all-encompassing” term referring to inputting technologies [scanning, indexing, optical character recognition (OCR), forms, digital creation, etc.], management technologies (document services, workflow and other work management tools) and storage (primarily non-alterable or write-once read-many) technologies. This Technical Report provides information to users related to what technical reports, guidelines and standards have been developed for technologies commonly available in document management systems.

Over the years, the industry has utilized various terms and acronyms to describe these core technologies, including, but not limited to,

— technology utilization,

— eCommerce,

— content management,

— B2B,

— P2G,

— G2G,

— knowledge management,

— EDMS,

— ECM,

— ERMS,

— EDRMS,

— EC3M,

— WCM,

— BPM,

— etc.

It is important for organizations to recognize that technology vendors commonly develop new terms and acronyms to present updated product technology and new uses for existing technologies. While these applications and/or products typically provide additional level of functionality, they are still based on at least one of the following core EDMS technologies, including

— document/library services,

— document imaging,

— forms management,

— routing/workflow, and

— ERM.

Additionally, it is important for organizations to recognize that many acronyms have different meanings, depending on which industry and/or organization is using those terms. For example, the acronym BPM is used to describe business process management, which is a process undertaken by the organization, is also used to

reference business process modelling, and is currently being used by some vendors and vendor-specific organizations to redefine workflow. The use of this term is a good example of how vendors re-use terms commonly utilized by the industry for other purposes. Throughout this Technical Report and specifically in the document subclauses describing various implementation process and activities, the business reviews how processes function and how the organization manages these business processes. From that perspective, the entire lifecycle of any EDMS project can be referred to as business process management. This is not to indicate that there can be only one definition for any term, but it is necessary for organizations to consider carefully the context in which the vendors/suppliers are using these redefined terms to ensure the desired/anticipated technology is implemented.

Another example is the use of ERM, which is used to describe electronic report management, but is also used by records managers to describe electronic record management. This Technical Report provides information related to those terms and acronyms recognized by the document management industry that best describe the underlying technologies, enabling readers to have a foundation from which they can determine what is required by the organization, regardless of the product name or acronym used by various vendors.

This Technical Report is not intended to be an all-inclusive paper on electronic document or content management and does not attempt to influence any single technology or provide legal guidance or legal opinions. While there are storage technologies other than optical/magnetic currently available (i.e. microfilm, microfiche and hybrid storage systems) that are not included in this Technical Report, it is advisable to review those technologies if they are determined to be appropriate by the end-user organization.

Document management — Analysis, selection and implementation of electronic document management systems (EDMS)

1 Scope

This Technical Report presents a recommended set of procedures and activities that are advisable when performing analysis, selection and implementation of project phases associated with electronic document management systems technologies. This Technical Report provides user-level information outlining specific recommended activities to be completed throughout the various project phases typically performed when implementing these technologies. It outlines the steps and activities, together with compliance with relevant industry standards and guidelines that need to be examined and "certified" to ensure relevant technologies have been analysed, designed, implemented and managed, ensuring document/record validity when used in a business or government environment.

This Technical Report provides both user- and technical-level information and guidance detailing specific recommended activities and project tasks/phases recognized throughout the EDMS industry as being the EDMS industry best practice related to analysing business processes, evaluating appropriate/relevant technologies and ensuring complete technology implementation where required by the organization.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 12651:1999, *Electronic imaging — Vocabulary*

ISO 15489-1, *Information and documentation — Records management — Part 1: General*

ISO/TR 15489-2, *Information and documentation — Records management — Part 2: Guidelines*

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 12651, ISO 15489-1 and ISO/TR 15489-2 apply.

3.2 Abbreviated terms

API	application programming interface
ASAP	asynchronous service access protocol
B2B	business-to-business
BPM	business process management

BPR	business process re-engineering
COLD	computer output to laser disk
COM	computer output microfilm
COTS	commercially available off the shelf
CSF	critical success factors
CWAD	common workflow audit data
DMA	Document Management Alliance
ECM	enterprise content management[1] or electronic content management[2]
EC3M	enterprise content, collaboration and communications management
EDMS	electronic document management systems[3]
EDRMS	electronic document records management systems
ERM	enterprise report management
ERMS	electronic records management systems
G2G	government-to-government
ICR	intelligent character recognition
IDL	interface design language
IT	information technology
OCR	optical character recognition
ODBC	open database connectivity
ODMA	Open Document Management API
OLE	object linking and embedding
OMG	object management group
OMR	optical mark reader
P2G	public-to-government
RFP	request for proposal
RMA	records management applications
SOAP	simple object access protocol
WAPI	workflow application programming interfaces
WCM	web content management
WfMC	Workflow Management Coalition

1) Enterprise content management is defined in ISO 12651-1 as a set of tools and methods that allows an organization to obtain, organize, store and deliver information crucial to its operation. It can be broken down into five major components consisting of capture, manage, store, preserve and deliver content.

2) Electronic content management is the same as EDMS in that it focuses on the technology aspects of the overall environment.

3) Although there is a difference between enterprise content management, electronic content management and electronic document management systems, in this Technical Report the acronyms EDMS and ECM are used synonymously from the perspective that both require the use of core technologies, together with policies, procedures and methodologies to successfully design, implement and manage electronically stored information.

4 Electronic document management technologies

4.1 General

Even in today's world, many organizations still function almost entirely in a “paper-driven” environment. This environment is a direct result of the need to maintain information on all aspects of the organization and can be seen throughout many organizations. When considering EDMS technologies, organizations should consider implementing the necessary foundational components and then add other functionalities as required.

Electronic document management systems (EDMS) has become an all-encompassing term, referring to the integration of various underlying technologies including

— document imaging (used to convert hard copy documents into digital format),

— document/library services (used to manage digitally born documents) (most EDMS systems allow users to use this technology to also manage scanned documents, if desired),

— workflow (used to route, track, and otherwise manage electronic documents and work activities),

— enterprise report management (ERM) (used to store electronic formatted reports),

— forms management (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.

Electronic document management systems provide users with greater access to digital information from a common user interface, through the utilization of industry standard Internet browser technology. One of the primary reasons users prefer this level of technology is the distributed functionality and extent of digital information availability that can be accessed almost immediately after implementation.

The structure of EDMS technologies can be viewed as a set of building blocks as noted below in Figure 1. 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 have resulted in the database components to be considered a “commodity item”, rather than a specialized tool.

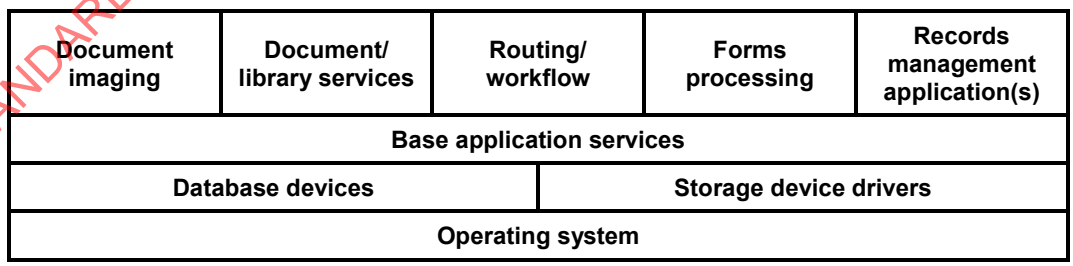

Figure 1 — EDMS technology building blocks

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 that integrate the core applications components with the database services and provide the components integrating the storage environment with the overall solution.

The fourth layer incorporates the various core technologies of EDMS. Each of these core technologies (with the exception of the operating system layer) are further described in 4.2 to 4.11.

4.2 Database services

There has been a significant shift from developing custom technology solutions at the database level to configuring or implementing commercially available software over the past 10 years. As the EDMS industry and associated technologies matured, end-user organizations were able to shift from a “development” model to a “configuration” model for the base technological components.

This is an important consideration for any organization evaluating EDMS technologies from the perspective that almost all of today's EDMS solutions have moved away from the need to have specialized database administrators towards actually discouraging organizations from changing and/or modifying the EDMS database table structures and configuration, which in many cases now result in the solution provider withdrawing solution support. While years ago it was important for the organization to hire dedicated EDMS database administrators, this is no longer the case. Over the past 5 years, the industry has noticed that almost all EDMS solutions (with the exception of highly specialized solutions) have effectively standardized on the most commonly used database platforms, enabling the solution providers to offer standardized support, thereby significantly reducing the organizational cost associated with hiring database specialists solely or primarily dedicated to managing the EDMS solutions.

4.3 EDMS application services

The next layer in the “building block” is considered to be the EDMS server application. Early in the development and maturity cycle of EDMS technologies, end-user organizations were required to provide database administration and resources. During the late 1980s and early 1990s, the EDMS technologies had not matured to the level approaching the commercially available off the shelf (COTS). While the EDMS technologies were maturing, end-user organizations were required to maintain the database together with the application.

Current versions of enterprise EDMS solutions have shifted the database administrative functions back to the vendor with the end-user organization responsible for daily application maintenance and periodic server maintenance. Most enterprise EDMS solution providers provide their products with technical support including system installation, initial configuration, application updates/patches, etc. This major shift from requiring significant technical resources at the end-user level to vendor-supported solutions has resulted in an extensive amount of vertical market penetration. This has been achieved through the use of the standard technology components configured to address specific environments and business needs.

4.4 Core technologies and application-specific modules

There are various core technologies and application-specific modules that provide specific functionality including some or all of the following core technologies: document imaging, document/library services, workflow, forms processing, etc. All enterprise EDMS solutions have at least one of these core technologies and in most cases include multiple components. Furthermore, many solutions have integrated both the document imaging and document or library services components into a single application, while other solutions simply integrate these components together as required by the organization.

4.5 Document imaging technologies

Document imaging technologies enable users to scan hard-copy 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. The following are the four basic components to document imaging systems:

— input;

— identification;

— storage;

— 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* stations allow users to identify (or index) incoming documents, allowing them to be retrieved at a later date. The *storage* part of the system consists of various components connected to the document management or workflow server and is 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.

4.6 Document or library services technologies

Document or library services technologies enable organizations to manage digitally born documents. Document or 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, together with a mechanism to store or manage digitally born document libraries. The basic capabilities of these technologies include allowing authorized users to

a) load or import digitally born documents directly into the system,

b) enter relevant metadata associated with the document,

c) create virtual folders linking various documents together,

d) check information in/out of the repository,

e) co-author documents also referred to as collaborative development,

f) secure access at document level,

g) maintain detailed audit trails,

h) make changes and check the modified information back into the repository,

i) manage whether original documents are updated or replaced during the update operations,

j) establish security levels for groupings of documents.

The management portion of document or library services technologies includes the ability to restrict access to certain documents or group of documents to only authorized users. Together 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 not be allowed to check it out, and they 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 records management versus content management versus archival management, but the underlying functions remain the same with slightly different terminology and/or descriptions):

— import;

— identification;

— storage;

— 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* components allow users to identify (or index) this digital information, allowing it to be retrieved at a later date, as well as providing a vehicle to store information related to the digital data itself [such as author, purpose, subject(s)], 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.

4.7 Workflow technologies

Workflow technologies can provide different levels of routing, tracking and administration. These technologies can be grouped into three categories:

— administrative;

— ad-hoc;

— 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 and provides extensive tracking and logging capabilities. When considering production workflow technologies, the organization should consider whether to use role-based or user-based 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 enabling users to control the process logic. 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 primary 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 tool in the automation of processes and information posting to a website. 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 various coalitions, standards committees and EDMS vendors have completed development of various standards.

The maturity of workflow technology and the associated trends are based on the separation of the processing rules from the processing scripts or work routing. In more sophisticated workflow environments, workflow scripts are tightly integrated to specific activities making the routing, editing, approval and submissions of content manageable at the user level. Interaction with the various thin-clients would trigger sub-processes as defined in the workflow script, resulting in the appropriate applet being downloaded and/or launched.

Workflow computing is the automation of work processes performed throughout an organization. A workflow application automates the sequence of actions, activities or tasks used to run the process. This includes tracking the status of each occurrence of the process and providing tools to manage the process. The following are the four basic components to a workflow system.

a) **Processes**: An automated workflow application is made up of the different tasks or activities that are necessary to be completed to achieve a business goal. The workflow engine manages these processes. The workflow application works in conjunction with the engine to manage the work process.

b) **Work queues**: Work items are created and distributed according to preset rules and placed into work queues. Users or groups of users are assigned to various work queues as required for processing. Work items within these queues can also be automated.

c) **Tools**: There are various tools accessed by the user, including forms display, word processors, terminal emulators, legacy applications, etc. These tools are used to access existing host applications and perform office related activities as required to complete work.

d) **Object data**: "Object data" is another term for any digital content referenced and/or used by the workflow system. The term became more prevalent after the computing technology was able to support video, audio and other forms of information into the workflow system. These objects become the work item to be processed during the normal course of business.

4.8 ERM technologies

Enterprise report management (ERM), which was previously known as computer output to laser disk (COLD), is an integrated software and hardware solution that stores and indexes formatted computer output (pages) on optical disks, magnetic disks, 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 with which ERM technology is concerned 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 COM. The structure and format of this output is known. This data is time-period focused, i.e. 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, lies a myriad of complex tasks. Data should be downloaded or transferred to the ERM system server before being 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 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.

4.9 Forms processing

The creation and utilization of electronic forms enables organizations to collect data in a standardized format and automatically enter or load the data into an EDMS solution. Electronic forms are typically created either using a forms design package or through the use of standard HTML editors. Forms design packages typically include not only the forms design components, but also enable organizations to "tag" or identify each field on the form and relate that data to a database or application that would receive and further process the information. These forms management tools also enable organizations to validate and/or perform edit checks on the forms as they are being completed to simplify data entry.

The use of forms within the EDM industry has become widespread and most EDMS solutions incorporate some level of forms design and/or management as a portion of the standard product offering. In many cases, the use of forms design and management tools are replacing the older style of programmed forms that was required in the 1990s. With using these tools, organizations are able to quickly develop and deploy forms driven data entry across the Internet without significant development efforts.

4.10 Optical and intelligent character recognition

Optical character and intelligent character recognition enable organizations to quickly capture information from hard copy documents that need to be processed after document imaging and storage. Optical character recognition (OCR) and intelligent character recognition (ICR) can greatly reduce the time required to index documents while enabling organizations to develop in-depth full-text searchable databases.

The value of using these technologies is especially evident when organizations need to capture specific portions of documents that are consistent. To capture this information, the utilization of "zoning" allows specific portions of similar documents to be identified and information within that "zone" captured and further processed as required by the application. Throughout the EDMS industry, it has been found that the use of these technologies can also greatly improve the quality of information being indexed, while reducing the overall staffing requirements to perform the same functions manually.

4.11 Records management applications

Another common application that organizations implement is referred to as records management applications (RMA). These applications provide the configuration and management tools enabling organizations to implement rules associated with identifying what length of time specific types of information (regardless of whether referenced as a document, record, image, etc.) and what type of disposition should occur when the end of life has been reached. 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 applications will only provide an automated retention schedule, hence the integration to the underlying EDMS technologies.

RMAs are considered to be software used by an organization to manage its records. An 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 should have at least one core EDMS component. Without at least one 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, i.e. 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.)

5 EDMS guidelines and standards

5.1 General considerations

Industry guidelines and standards enable organizations to follow industry-accepted practices and procedures. Standards and recommended practices specified in a federal, state, provincial or local law/regulation 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 should 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. A list of guidelines and standards is given in Annex A.

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 detailed information necessary to successfully prepare for, select and implement the desired technology. The guidelines that users should evaluate include the following:

— request for proposal (RFP) guidelines;

— recommended document preparation procedures for scanning/indexing;

— planning considerations for technology implementation;

— how to determine what information should be included during document indexing;

— legal considerations;

— forms design;

— selecting the appropriate image compression methodology to be used;

— sampling procedures to verify information being scanned and indexed;

— establishing quality requirements and quality control.

The industry standards include standards related to document services integration and toolkits, workflow integration and toolkits, document imaging related standards, and optical storage standards. Product suppliers should certify that their products meet the specified standard(s) to ensure that the product is, in fact, compliant with the relevant standard(s). It is important to note that as the industry creates and approves new standards and guidelines, this Technical Report will be updated to reflect those changes after the standards/guidelines have completed ISO approval processes.

5.2 Trusted EDMS technologies

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 or 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/TR 15801 and those related to records management policies contained in ISO 15489-1 should be considered. This will ensure that both technical planning, design and implementation, together 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 should be maintainable in an ongoing fashion after rollout into production.

As described in detail in ISO/TR 15801, 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 original format. The trusted document management system should ensure that at least two (2) separate copies of the electronically stored information are created on electronic media and that they at a minimum meet all the following conditions.

a) The trusted document management system should utilize both hardware and media storage methodologies to prevent additions, modifications, or deletion of information to the original document or record during the approved lifecycle of the stored information.

b) Hardware and media storage methodologies used to store information in a trusted system should be verifiable through independent audit processes.

c) The trusted document management system should 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.

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. (Additional information related to all aspects of the trusted system are documented in ISO/TR 15801.)

5.3 Industry guidelines

5.3.1 Review

5.3.1.1 General

Industry guidelines should be reviewed and will assist the organization during the preparation, planning and implementation phases of the document management project.

5.3.1.2 Terminology

To gain an understanding of various terms used throughout the industry, the organization should review ISO 12651. This Technical Report provides a detailed list of various terms that will be encountered during discussions with product vendors and integrators.

5.3.1.3 Human and organizational issues

When implementing these technologies, the organization will face several human and organizational issues. ISO/TR 14105:2001, about human and organizational issues for successful implementations of electronic image management (EIM) systems, provides detailed information gathered throughout the industry related to system usability and adoption by the users. These guidelines will assist the organization during all the change management activities required for successful system implementation.

5.3.1.4 Request for proposal (RFP) development

Prior to selecting a specific product/integrator, the organization should document system requirements and provide them to the vendor(s) or integrator(s) being considered. Regardless of whether the RFP is being sent to a single vendor/integrator or multiple vendors/integrators, an RFP should be developed to enable the organization to clearly define their requirements and enable the vendor/integrator to clearly understand all business and technical goals and requirements.

5.3.2 Legality issues

5.3.2.1 General

Legality issues that should be considered by legal advisors include information expungement, retention requirements and information redaction. Each of these issues is organizational dependent and should be considered by legal advisors. Organizations should implement trusted EDMS solutions when it is important to store and manage authentic documents or records that maintain their reliability and integrity over time, as well as ensuring usability/readability. Additional information related to all aspects of the trusted system is given in ISO/TR 15801.

5.3.2.2 Expungement

Expunging information from databases and storage systems needs to meet specific legal requirements (see ISO/TR 12037). 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.

5.3.2.3 Retention requirements

Users and system designers should consult the organization's established retention requirement set forth in their record management policies and procedures (see ISO 15489-1). 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 should be considered, hardware and software obsolescence issues should be evaluated, and a sound migration strategy should 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, as well as provide a mechanism to automate when documents are to be archived, for how long, what action to take after the retention period is passed, together with numerous other organizational advantages from a management perspective.

5.3.2.4 Redaction

The process of redaction (see note) 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.

NOTE “Redaction” refers to a process by which parts of a document are kept from disclosure. For example, the name of a person, a personal identification number, or entire paragraphs that reveal trade secrets. 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. 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.

5.3.3 Technology standards

Technology standards are developed for specific technologies and not at the “general” level. All industry-related and relevant standards are listed in the appropriate technology category within this Technical Report.

5.3.4 Implementation considerations

5.3.4.1 System administration

When selecting the technologies to support the business requirements, the organization should receive and maintain detailed information related to system administration functions required to administer and control all applications, security, system server hardware and data backup/migration. The product supplier or system integrator should provide this information. These requirements should include:

— operating system management (updates, patches, backup, restore, etc.),

— application software (updates, patches, backup, restore, etc.),

— system security (user additions/deletions, security modifications, etc.),

— data migration (retention periods, media replacement, etc.),

— software trouble-shooting tools,

— hardware trouble-shooting tools, and

— database management utilities.

These technologies can be hosted off-site as well as on-site. The information technology (IT) group within the organization should be provided the tools to perform these system administrative functions. At a minimum, the product supplier or system integrator should allow the IT group to manage the system and utilize the technical/support staff within the product supplier to resolve application and/or database issues that may be encountered as well as assisting in software updates.

5.3.4.2 Security requirements

To ensure the technology supports secure access that meets the organization's business needs, the solution should also be assessed with respect to how it supports end-to-end security as related to user authentication, document authentication, and secure network transactions over the Internet, intranet and extranet, as necessary. Understanding the complexity and scope of an organization's security issues, especially when dealing with the Internet will require the collaboration of multiple organizational disciplines including legal, business operations, system administration, network administration, vendors and external users of the system. For more information on security related requirements, organizations should review ISO/IEC 17799.

5.3.4.3 Capacity planning

To ensure an accurate assessment of the scope and size of the document management system, the organization should attempt to determine the capacity requirements of the expected solution. While the natural tendency of any organization is to regard all documents as essential and to store all documents with equal access, the selective categorization of documents by type, retention period and frequency of access may contribute greatly to the final cost of the solution. Categorization of documents based on an organization's usage statistics is vital to determining how the document is finally stored in the system, including optical library units and online storage. The following is a partial list of some sizing parameters to consider:

— system availability requirements;

— number of form types and documents;

— retention requirements by document or form type;

— frequency of document access;

— peak daily volume of new documents processed;

— volume of new cases for workflow consideration;

— number of internal users (case workers, researchers, data entry operators);

— number of users (local and remote);

— number of organizations;

— number of remote sites.

5.3.4.4 System performance

To ensure that information is available for use by the users within anticipated time frames, it is important that specific requirements related to performance expectations be defined. The organization should determine the anticipated response times they expect from the system for

— document retrieval from long-term storage media,

— document retrieval from online cache,

— document viewing (over the LAN, WAN, internet),

— document printing, and

— scanning/indexing performance.

5.3.4.5 System scalability

Organizations should ensure that the solution be scalable. This scalability includes the ability to increase the number of processors in a multi-processor environment, increase the number of servers to operate in a cooperative fashion, as well as increase the storage capabilities as required by the organization. Requirements should include

— the ability to increase the number of system users without component replacement,

— the ability to support other technologies, i.e. OCR, form management, etc.,

— the ability to support multiple servers and standardized non-alterable write-once storage solutions in a distributed manner, and

— the ability to support symmetrical multi-processing, if required by the organization.

5.3.4.6 Fax services

Facsimile (fax) transmission services enable users to send and receive faxed documents directly at their workstation. When considering these services, organizations should evaluate the following requirements to support their specific business needs.

— **Outgoing fax without document viewing**: This allows for users to fax documents directly from their computer without viewing each document first. The user should be able to select a range of documents and have them routed to the fax “server” for transmission.

— **Outgoing fax after document viewing**: This allows users to fax a document during viewing. The user should be able to attach other documents to the outgoing fax, if needed.

— **Incoming fax processing**: As incoming documents are received, the system should allow users to receive incoming documents and automatically route the document based on configurable rules (via a system administration interface) either by incoming telephone number or through forms or OCR processing.

— **Fax status reporting**: The system should provide a fax reporting capability enabling users to view status and historical information related to faxes sent by the user. This historical reporting should be based on user security rights, preventing users from accessing other users' history, while allowing users with higher levels of security to access all historical records.

5.4 Document imaging

5.4.1 User guidelines

5.4.1.1 General

User guidelines should be reviewed for document imaging technologies to assist the organization during all project phases from the planning phases through to the actual implementation. These guidelines will greatly simplify the overall process and ensure that critical aspects of preparation and implementation planning are addressed early in the project rather than forcing the project to stall until planning issues are addressed.

5.4.1.2 Planning

During the planning stages of the project, the organization will need to address various issues including planning for the implementation of document imaging and preparing the documents to be scanned.

5.4.1.3 Indexing the information

As the organization plans to implement document imaging, it should consider establishing relevant indexing field guidelines and procedures. These organizational guidelines should provide detailed information that should be considered when planning the indexing requirements for all current and anticipated documents to be scanned. Establishing all the necessary index values with the ability to add additional and/or other document types prior to system implementation greatly improves the value and quality of information being scanned and stored in the system.

5.4.1.4 Storage technologies

During the planning stages of the project, the organization should review guidelines for the planning, implementation and operation of long-term and permanent record storage on electronic media (see ISO/TR 15801). There are various approaches to document or record storage technologies, including using optical WORM for long-term preservation or magnetic WORM for those organizations who need faster retrieval speeds and are less concerned about long-term archival. Another approach is to use magnetic WORM for temporary storage cache (where users can retrieve documents quickly) and optical storage for long-term storage. Both sets of technologies are viable for document/record storage depending on user needs, regulatory storage requirements, etc. Users should exercise caution when using non-standardized or proprietary storage technologies.

5.4.1.5 Image formats

The organization should ensure that all information being scanned or electronically received is stored in an industry-accepted format such as TIFF, JPEG, JBIG, JPEG 2000 or PDF/A. Non-standard or proprietary file formats should not be used. Non-standard or proprietary formats include any formats used by a single vendor/source and accepted as a standard file format at either a national or international standards level.

Proprietary file formats include “file-wrappers” used to encapsulate standard file formats within a non-standard structure.

5.4.1.6 Indexing quality control

As the system is moved into production, it will become important for the organization to develop a methodology of reviewing both index data and the actual documents, to ensure the information is available and readable. The organization should establish a documented process to ensure that all documents are properly scanned and indexed. This documentation should be followed by all personnel performing scanning and indexing, and should provide a mechanism for index data entry verification prior to document committal to the storage media and/or transmission to the business process.

5.4.1.7 Scanning quality control

Scanning quality control measures enable users/operators to ensure that the scanner is operating within anticipated tolerances. ISO 12653 provides additional information for production document scanners. Following these procedures will enable the user/operator to ascertain whether the scanner is properly set up before scanning actual documents.

5.4.2 Technology standards

Technology standards in this area are currently being developed. As these standards are completed and approved, those applicable within the scope of EDMS technologies will be incorporated.

5.4.3 Implementation considerations

5.4.3.1 Document scanning

The document scanning part of the system should provide the ability for the users to quickly digitize documents and route these documents to the person performing the indexing operation. Requirements associated with this part of the system should include the following:

— the ability to support both batch processing and single document scanning and indexing;

— the ability to support document re-scanning;

— the ability to support both simplex and duplex scanning;

— the capability of the scanner to scan at the resolution meeting the specific image quality requirements of the system, such as 200 DPI, 300 DPI or 400 DPI; and

— the ability to set page breaks when batch scanning fixed and variable length documents.

5.4.3.2 Document scanning and indexing

When implementing document scanning and indexing technologies, the requirements should include detailed information related to all processing phases. If colour documents are to be scanned so that the image captures the colour, the scanner should be capable of doing so. Patch code and bar code hardware and software should be included if these techniques are to be used for the automation of data indexing. When using these technologies, the user should be aware that bar coding and OCR technologies typically minimize key stroking during the indexing phase but do not always negate the need for manual indexing. The level of information captured automatically will vary depending on the quality of the incoming document and the ability of the system to accurately recognize the required information.

The issue of performance is of critical importance and the organization should ensure that the selected solution provides the ability to scan and index documents within anticipated time frames. The various processes associated with document scanning and indexing includes the following:

— the time required to prepare the document for scanning;

— the scanning of the documents, ensuring all documents and all sides (for double-sided documents) are captured;

— the time required to index and verify the documents;

— the time required to route the document to the end user for further processing (if workflow technologies are being used);

— the ability to preset common fields (for indexing purposes) when scanning in batch mode; and

— the ability to support auto-indexing of documents using bar codes, OCR or ICR.

5.4.3.3 Scanning or indexing throughput

The system should be capable of scanning either single or double-sided documents. Scanners should be capable of processing the daily work volume at the selected scanning site. This processing will include document preparation, scanning and indexing. The system should also be capable of supporting low, medium and/or high volume scanning capabilities depending on user requirements and selected scanner. The total number of scan stations and indexing stations should be determined by the organization to ensure that all work can be processed within anticipated time frames and stations are available for use when needed.

5.4.3.4 Document image compression

Image compression/decompression should support ITU Group 4, JPEG, JPEG 2000, JBIG, or other output format standards with no proprietary alterations of the algorithms. The selected compression technology should not include extraneous information unsupported by relevant industry standards. Users should 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 that are available. ISO/TS 12033 is a specification that provides information enabling users to select the appropriate compression technology which the vendor/integrator should support for different types of data. The different types of data may include scanned documents, line art, photographs, etc.

5.4.3.5 Post-scanning processing

Post processing may be used to provide image “clean-up” after the scanning and prior to indexing and final storage. This software generally performs de-speckling, de-skewing and other functions to improve the quality of the scanned image with limited operator intervention.

Image “clean-up” and other post-scanning processing should only be used to improve legibility. Caution should be exercised when using these tools, as any material modification to the image may affect the ability to authenticate the document in a legal proceeding.

5.4.3.6 Document indexing

Detailed information related to all aspects of document indexing should be clearly defined. This should include the ability for users to index documents on workstations other than the scanning stations and support the ability to:

— index images either prior to storage or immediately after storage,

— identify both mandatory and optional indexes,

— add other indexing values, and

— batch index.

5.4.3.7 Optical mark reader (OMR), OCR, bar code and ICR processing

The main objective of the available recognition technologies is to reduce the amount of manual data entry for the capture of both hand-printed and machine-printed information from digitized documents. Although the technology will never eliminate the need for manual data entry, the effective use of these technologies on targeted documents has produced remarkable benefits often evaluated in reduction of manual keystrokes. The following is a brief list of evaluation criteria to consider when analysing the use of automated data capture.

— Is it possible to identify documents with sufficient volume to justify automated data capture processing? These are typically used with forms containing both structured and unstructured content, and with identifiable information to be extracted.

— Is it cost effective? Determine the amount of data to be captured and the cost to support a manual solution, and then compare it to an automated data capture solution.

— Is it possible to re-design the target forms for improved recognition? The use of checkboxes, patch codes, bar codes, dropout ink and OCR fonts all provide considerable improvement in recognition accuracy rates.

— How will the documents be batched for scanning? Will it use mixed form sizes? Will the scanner accept mixed form types? Is it possible to introduce a batch header sheet to streamline the scanning process?

— Identify the business rules that may be used for post-recognition processing to improve the accuracy of the information captured. For example, the capture of a unique personal identifier can be used to automatically verify the name and address information against the organization's existing database.

5.4.3.8 Quality control

When defining quality control for document scanning and indexing, the organization should include the ability for the user to be able to:

— check and validate the complete scanning and indexing process,

— facilitate the re-scanning of poor quality images,

— verify readability of each page of each document,

— verify proper indexing of each document,

— verify accurate page counts for each document, and

— verify accurate security for each document.

5.4.3.9 Query and retrieval display time

Query and retrieval display time is commonly of high importance to the users. The user should define the anticipated performance requirements prior to system design and hardware procurement. These performance requirements should include maximum response times anticipated during production. This should also take into account the total number of anticipated simultaneous user requests; the total number of drives; whether the information is available in an on-line, near-on-line or off-line mode, etc.

These time periods include all time required to retrieve the appropriate optical/removable media (when necessary), reading all requested pages from storage media, storage of all requested pages on magnetic cache (if being used) and subsequent transmission of the first page to the user for viewing. When removable media (i.e. optical WORM, CD, DVD, tape, etc.) is implemented, this response time should take into account time required to: "spin" the drive down, eject the media, retrieve new media from the storage bays, insert the media into the drive, "spin" the drive up and retrieve information from the media.

5.4.3.10 Printing times

The imaging system should be capable of printing user-selected documents within anticipated user-established time frames. This response time includes document retrieval from optical storage and transmission to the selected printer. The user should have the ability to select a document, or range of documents, to be printed, without being forced to view any of the pages prior to print submission.

5.5 Document/library services

5.5.1 Technology standards

5.5.1.1 General

Technology standards should be evaluated by the organization to determine which standards are important and relevant to the overall project goals and objectives.

5.5.1.2 Open source distribution

The product vendor or supplier should certify that the organization can use open source document services software and metadata definitions (information describing the document) with their specific product. This will enable the organization to integrate other document services technologies without significant system re-development.

5.5.1.3 Development toolkits

The product/vendor supplier should certify that the system uses industry standard application programming interfaces. This will enable the organization to implement a document services system and access information stored on other document services implemented throughout the network. These toolkits simplify application development and will enable the organizations to develop a common user interface regardless of the product used to "house" the actual data.

5.5.2 Implementation considerations

5.5.2.1 General

Document services enable users to create, modify and manage electronic files typically associated with various office processing applications. These capabilities are included.

5.5.2.2 Version control/check-in and check-out

The organization should ensure that the product fully supports version control and check-in/out methodologies. Version control should automatically update the version number when a previously "checked-out" document is

returned to the information repository. The system should prevent more than one person from checking documents out for modification and use a security model ensuring that only authorized personnel can perform these functions.

5.5.2.3 Logical folders

The ability for the users to "logically" link a single document to multiple folders is important to prevent document duplication. The organization should ensure that the selected product allows an authorized user to create a copy of a document within a specific folder, or set of folder(s), while maintaining only one physical copy of the document within the system. The system should provide information related to which folders are "linked" through a query mechanism available to authorized users.

5.5.2.4 Group/user security

The system should allow organizations to apply security access/restrictions at both group and user levels. Group level security should apply to all users within the defined group, while user level security should provide additional security restrictions or capabilities for specified users beyond that established for assigned groups.

5.5.2.5 Document security

The system should allow organizations to apply security at the document or file level. Only those users with appropriate security levels should have access to these documents and/or files. This security should include read, update, annotation, highlighting, "mark-up" and creation control.

5.5.2.6 PDF/A, HTML, XML conversion

The system should provide for PDF/A, HTML or XML data conversion as required by the organization. This conversion should enable the users to convert existing office documents into a standardized format that can be accessed through a standard web browser.

5.5.2.7 Document publishing to a website

The system should allow an organization to update an existing web page automatically after completion of a review/approval process, or manual review and convert by the "webmaster". This document publishing functionality should include the ability to store native file formats or utilize web templates to reformat the document into either HTML, XML or PDF/A format.

5.6 Workflow

5.6.1 Technology standards

5.6.1.1 General

5.6.1.1.1 Workflow reference model

Technology standards have been developed by the Workflow Management Coalition (WfMC) into a workflow reference model. The significant aspects of the workflow reference model can be summarized into the following three categories, each building incrementally on the preceding.

a) A common vocabulary for describing the business process and various aspects of the supporting technologies to facilitate automation. This provided the essential foundation for the subsequent detailed discussion on how a workflow system could be architected in a general sense.

b) A functional description of the necessary key software components in a workflow management system and how they would interact. This was developed in a "technology neutral" manner, to allow the model to be independent of any particular product architecture and implementation technology.

c) The definition, in functional (or abstract) terms, of the interface between various key software components that would facilitate exchange of information in a standardized way, thus enabling interoperability between different products. Five such interfaces were identified and became the foundation for the WfMC standardization programme.

It was an important principle that the reference model focussed specifically on workflow management technology and standards. It deliberately did not attempt to define standards in other related areas in which other industry bodies were working; these were seen as complementary.

5.6.1.1.2 Five interfaces

Each interface was initially specified as a business level statement of objective, in order to identify what the interface was intended to achieve in business terms and why a standardized approach was desirable. This was subsequently followed by a detailed but abstract specification of how the interface operated and finally (for most interfaces) a "binding" specification covering the implementation of the interface in a particular technology.

The five interfaces are the following.

a) **Interface 1** (i/f 1) was developed to support the exchange of process definition data between BPR tools, workflow systems and process definition repositories, enabling users to select the most appropriate tool for different aspects of the business process lifecycle. It was specified as a process definition meta-model, defining the process objects, their attributes and relationships, and a textual grammar for expressing the process definition structure and information content. This was subsequently re-expressed as an XML document definition (XPDL).

b) **Interface 2** (i/f 2) was developed to facilitate client application integration with different workflow systems, in particular to support the principle of (client) application portability and reuse with different workflow management systems. It was specified as a series of workflow APIs (WAPI) to allow the control of process, activity and worklist handling functions. These were originally defined in "C" and subsequently re-expressed in IDL (as part of the OMG workflow management facility) and OLE. A set of "C" APIs for manipulating process definition objects and attributes was also defined.

c) **Interface 3** (i/f 3) was developed to provide a common framework for third parties to integrate other industry applications and services, including specific support of agent interfaces to provide a common framework for access to legacy applications. It was developed as a set of five basic API calls, defined within the WAPI document to support a common mechanism for connection, disconnection and calling to a variety of agents or other third party software environments.

d) **Interface 4** (i/f 4) was developed to facilitate process automation across multiple heterogeneous implementation environments. It comprises an interchange protocol covering five basic operations, specified in abstract terms (initially it was defined in IDL) and with separate concrete bindings. The initial version was defined as a MIME body part for use with email; subsequent versions have been specified in XML (Wf-XML). Ongoing work has lead to version 2 of Wf-XML, layered over SOAP and ASAP.

e) **Interface 5** (i/f 5) enables consistent audit and administration of workflow cases across systems, through the specification of a common model for audit data, including event identification, formats and recording. As such, it was originally specified in abstract terms, although a set of common APIs for access to audit data was subsequently developed. Recent work is aimed at expressing the audit data structure as a set of XML structures.

Although conceived as five individual interfaces, the separation is apparent only when viewed in the context of the stated business objective. In reality, there is significant commonality of function between the various "interfaces"; for example, the triggering of the initiation of a process execution is fundamentally the same action whether it is done client side (i/f 2) or server side (i/f 4). The evolution of the WAPI (API) specification started with client application interactions but expanded to include a full repertoire of API calls. Similarly, Wf-XML was developed initially for server-server interaction but has also been used successfully for client-server interactions.

A more useful and fundamental distinction is perhaps to take a view of each interface from the perspective of process ownership and administration control. In particular, interfaces 2 and 3 may be considered to be "tightly bound" to the local workflow management system and reflect a local view of resource management, i.e. interface 2 handling interaction with human resource and interface 3 interaction with automation resource. This has two significant consequences.

In the first place, the process definition is localized to the point of process enactment through the expression of the resource assignments (e.g. participants and applications). Secondly, the reference model could make the simplifying assumption that specification of messaging between a WFMS and participants need not be contained in detail within the process definition. It becomes a function of the WFMS locally to organize the most appropriate form of interaction with the participants via local *Worklists* (web access, email, etc.), according to the defined (within the process definition) *Activity* or *Procedure.*

5.6.1.2 Workflow development toolkits (WfMC interface specifications 2 and 3)

The vendor should certify that the product supports workflow application programming interfaces (WAPI). These APIs, as described in Workflow Management Coalition (WfMC) documents, ensure the implemented product provides a consistent method to access workflow management functions, particularly in cross-product implementations.

5.6.1.3 Workflow auditing (WfMC interface specification 5)

The vendor should certify that the product supports the WfMC audit specification. This specification details information to be captured and managed by the workflow system during operation. This will ensure that all relevant data is associated with all functions within the workflow technology.

5.6.1.4 Workflow interoperability (WfMC interface specification 1)

The vendor should certify that the product supports industry interoperability standards including the usage of standard e-mail systems. These interoperability standards will enable the organization to share workflow information directly between different workflow systems without requiring specialized development.

5.6.2 Implementation considerations

5.6.2.1 General

Implementation considerations should include the elements described in 5.6.2.2 to 5.6.2.9.

5.6.2.2 Workflow

Workflow technologies include various types of routing including ad-hoc routing, administrative routing and production routing. Ad-hoc routing enables the user to specify a specific process for a document to follow for that document only. Administrative routing enables users to define specific routing for a specific type of work that is always followed, regardless of the data within the work being routed. Production routing enables the users to define rules and work methods based on the document type and data contained within the work item. As the data changes, the production routing system would process the document accordingly, including the ability to support work timeouts, escalation and work reassignment.

5.6.2.3 Role versus user

There are two approaches to defining users within a workflow environment. The first method is to define a specific user to manage a specific task or activity. The second approach is to define a role within the work task or activity and then assign as many users as necessary or appropriate. The organizations should require a role-based system when implementing production workflow technologies.

5.6.2.4 Routing requirements

For those organizations requiring production workflow, the system should allow a user to route a document to another user. The following capabilities should be considered:

— the ability to automatically route documents into a routing queue based on document type or “type of work”;

— the ability to support multiple routing queues for each user based on the “type of work”;

— the ability to sort/retrieve documents in a routing queue in date order;

— the ability to sort/retrieve sections in a routing queue in “type of work” order;

— the ability to sort/retrieve documents in a routing queue in document type order;

— the ability to sort/retrieve documents in a routing queue for a specific person;

— the ability to change a “pre-defined” routing path;

— the ability to “pend” or “hold” items in that user's routing queue for work at a later time;

— the ability to retrieve specified documents from the routing queue on demand;

— the ability to define which documents require additional documents prior to forwarding;

— the ability to define timeframes for when additional documents should be received;

— the ability to define action to take if specified documents are not received by a specified date;

— the ability to process defined documents as a “logical” folder.

5.6.2.5 Graphical “rule designer”

The system should allow authorized users to create and modify work rules associated with the workflow system. This should include graphical-based design and management tools that would be used to create/modify work rules within a Windows or browser-based user environment.

5.6.2.6 Work monitoring

When selecting workflow technologies, the organization should evaluate whether work monitoring is required for their operation. Work monitoring tools enable the users to monitor current ongoing work, in a real-time basis (typically). This work monitoring is used not only for “load-levelling” of ongoing work activities, but also to see if there are any “bottlenecks” in the overall workflow process.

5.6.2.7 Escalation procedures

For those organizations requiring production level workflow, the selected solution should include the ability to automatically route work to a different user based on a specific rule or set of rules. The solution should also include the ability for users to manually escalate work as appropriate. During this escalation procedure, the solution should have the ability to have the work item returned or permanently reassigned, as determined by the user.

5.6.2.8 Error handling

As workflow items can include information not previously anticipated during the rules definition, the organization should require that the solution include the ability to handle errors within the routing of work

through the workflow engine. The error handling should include the ability to pre-define a role that would receive the appropriate work items that are determined to be in error.

5.6.2.9 Time-out procedures

When workflow is implemented, there are many instances where the timeliness of completing a specific work activity, or group of activities, is important. The ability to establish timers for all work items becomes very important. The organization should require that the solution support “timer” mechanisms and that the user is able to set these time-out values for specific activities throughout the graphical work “rule designer” tool.

5.7 COLD/ERM

Industry standards and user guidelines in this area are currently being developed. As these standards and guidelines are completed and approved, those applicable will be incorporated.

6 Best practices associated with EDMS project phases/activities

6.1 General

The following steps and/or activities provide industry-accepted guidance to assist end-user organizations in all aspects of the EDMS project. These steps include (but should not necessarily be limited to) business/operational process review, business/operational/technical requirement documentation, evaluation and selection of appropriate technologies addressing specific business issues, conversion/data issues, records retention and management issues, and implementation activities. Each of the following subclauses provides detailed information on those activities requiring completion prior to product/vendor selection. There are numerous steps and procedures associated with analysing business requirements through the identification and selection of relevant technologies to be considered for implementation. Annex B provides a high-level checklist of best practices associated with EDMS project phases/activities and provides a high-level listing of recommended project steps/activities associated with a full EDMS project, from process analysis through to system/product implementation.

6.2 Process/procedure baselining

6.2.1 General

The purpose of process/procedure base lining is to clearly define existing processes/procedures and identify issues and problems currently encountered. This is achieved through a detailed analysis of existing processes and procedures. When performing this analysis, it is important to capture and document activities including the following:

— how documents and information are received;

— what occurs to these documents after receipt (i.e., stamping, sorting, logging, delivery, etc.);

— how these documents are used and how many people use the same document to complete a specific activity or process;

— what happens to the document during the processing (annotation, highlighting, copying, etc.);

— after the processing is completed, where the document is stored, whether there are multiple copies, etc.;

— how the established document retention timeframes are adhered to and the process of document destruction after reaching the destruction date within the retention policy.

This information should be gathered through interviews with selected users within each processing unit. These users should include experienced users (non-management) and management personnel. It is important to

note that the team gathering this information should represent the business units from a user perspective and include all processes and procedures currently being used. As the base lining process continues, users may describe processes and/or procedures that are not “officially sanctioned” in the day-to-day processing. These workaround or alternative methods need to be documented, as well as all other user workarounds and methodologies implemented to complete daily work activities.

Upon completion of this documentation, the users should have an opportunity to review the baseline document to ensure that all functions and activities related to their processing have been accurately captured and documented. It is very common for these documents to have multiple versions presented prior to user sign-off. This is due to the primary fact that most users do not have complete documentation at the detail level related to how the documents are managed.

There are three basic activities recommended throughout the industry related to process baselining. These activities include the development of a high-level baseline establishing the overall structure of the business process, the detailed baseline documenting specifics of each task/procedure identified during the high-level baselining and associated manual processing metrics.

6.2.2 High-level baseline

The first step in documenting business processes is to develop a high-level (e.g. management view) of the manual- or document-based business processes. Selected and representative users from organizational staff should be interviewed to identify all general work activities, policies and business procedures. These processes should be documented in a “graphical” format developing the “high-level” process schematics documenting the manual- or document-based processing flow throughout the organizations. These “high-level” schematics are then further “exploded” during the detailed process documentation.

6.2.3 Detailed process baselining

Upon completion of the high-level processing schematics, those areas identified by managers and supervisors are further examined. Selected users are further interviewed from the detailed processing perspective. These interviews should include discussing how work and work-related information is received, processed, and “moved” between groups, departments and other users.

These processing schematics should be documented using a graphical tool enabling the organization to review the processes in an interactive fashion on their computers. A read-only version of the software should be provided to any organizational resource with the need to review this documentation. Users both identified during the initial project planning stages and others identified throughout the process should be interviewed to ensure that all relevant portions of the business processes are documented, showing all hand-offs and any other work-related activity.

During the detailed process baselining activity, processing rules and conditions are identified as decision points and document routing/hand-offs as the document moves through each identified process. “Rules and conditions” are considered to be those decision points and hand-offs that dictate how information flows through the process. A small example showing the level of information gathered detail for each process including “rules and conditions” is shown in Figure 2.

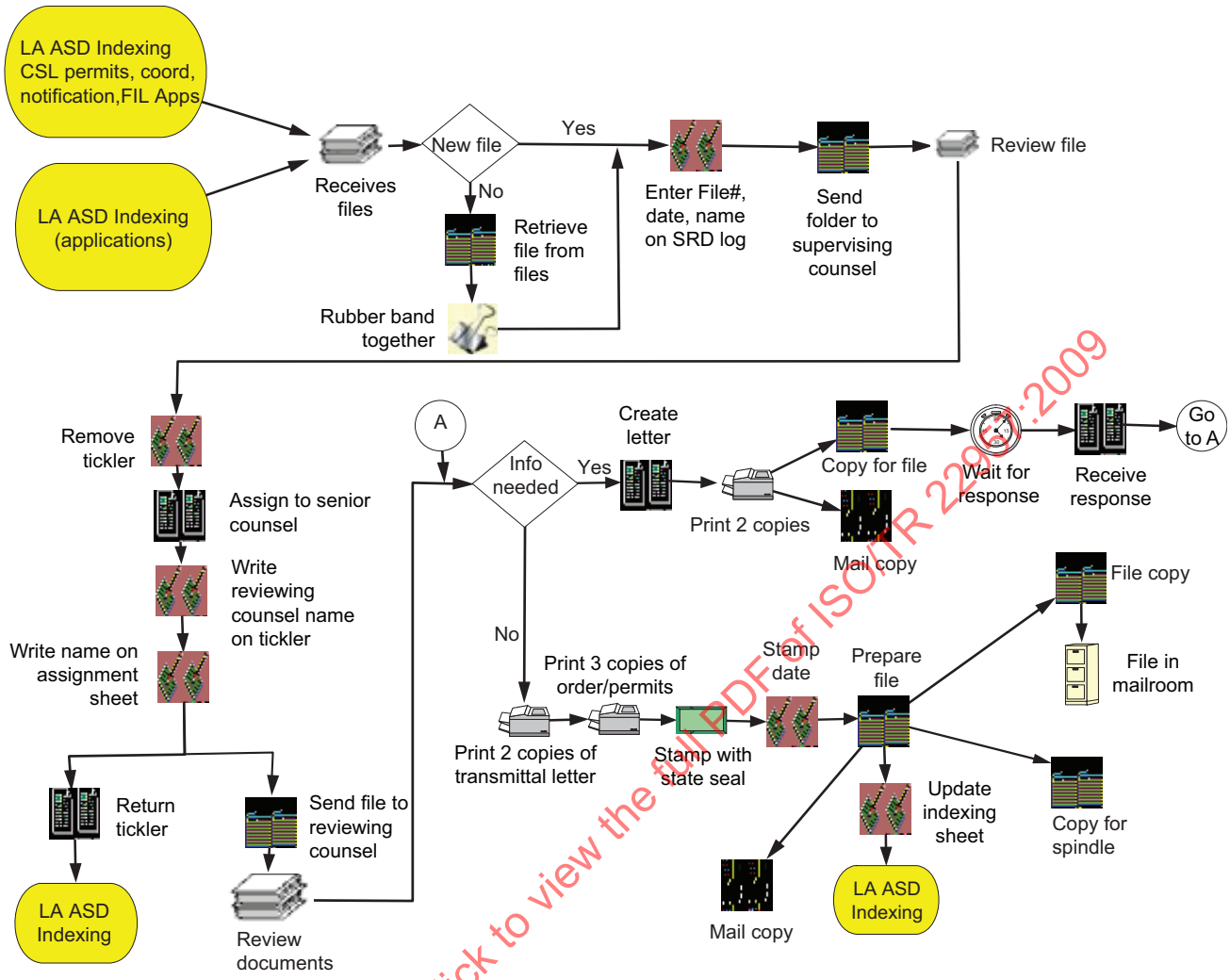

Figure 2 — Sample detailed processing schematic

6.2.4 Processing metrics

A processing metrics report that includes information related to the time personnel currently spend on all work related "manual" work activities, identified through the high-level and detailed schematic development, should be developed. These tasks can then be further examined by organization management to identify those processes and/or activities that can potentially be replaced/enhanced with technology and those processes/activities that can be replaced/enhanced through organizational change management.

This enables the organization to evaluate how much time is spent managing the current workload together with anticipated time after new/updated technology implementation. Time associated with processes/activities documented in the schematics should be identified and documented to identify items similar to the following:

— time spent logging receipt of documents;

— time spent copying, filing, locating documents;

— time spent manually managing digital documents from creation through to storage;

— time spent manually routing and tracking documents as they are processed;

— number of personnel performing associated major processes/activities;

— classification of documents and associated document volumes;

— estimated number of multiple copies of identical documents throughout the organization.

As accurate time related information is typically difficult for users to gather, it is common for an organization to follow a very conservative approach, evaluating how much “work” time is spent on these activities on a daily or weekly basis and then reviewing the information prior to inclusion in this documentation, while comparing the received information with overall times for the processes and other aspects of organizational work related activities.

6.3 Anticipated processes/procedures

Upon completion of the base lining process, this information is evaluated to determine where non-technology-based and technology-based changes could be implemented.

Examples of non-technology-based change typically include reduction in document copies, the revision of outdated procedures, elimination of redundant procedures and duplication of processes/procedures between organizations. During the non-technology change review, end-user organizations should consider the impact on existing operations when updating/changing processes/procedures.

Examples of technology-based change typically include automated logging of document receipt, automated routing for processing and detailed history related to work activities associated with each work item or document. Additional examples of technology-based change are related to which technologies are determined to be of benefit to the organization and how those technologies would be implemented resulting in a different method of conducting business.

The level of detail associated with the new or anticipated processes should take into account how legacy systems would be updated, modified and/or replaced. It is very typical to have a high number of personal databases, spreadsheets and other tracking/data tracking activities that can be consolidated into the EDMS system. It is also very common to have new processes and procedures established to not only manage digitally born information, but to support electronic distribution of this information to users within the organization and to users external to the organization through the use of electronic transmission.

6.4 Technology requirements definition

After identifying the relevant technology-based changes required by the organization, solution requirements should be documented providing detailed information to potential solution vendors. This document should clearly define anticipated user and system functionality in sufficient detail, enabling potential solution vendors to understand the business problem/issue being addressed and desired results after solution implementation.

When developing the solution requirements, organizations should consider documenting desired and/or required document management functionality and capabilities identified throughout this document.

6.5 Document classification and indexing model

An important aspect of any EDMS 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 EDMS solution would be used to discuss existing document classifications 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:

— be flexible and be able to be updated as required by the organization without adversely affecting documents already stored by the solution;

— fully address records retention issues and schedules;

— enable staff to easily locate and retrieve documents;

— support the ability to implement document security;

— provide logical information groupings;

— 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.

6.6 Business objectives and requirements

6.6.1 General

The business objectives, functional requirements and expectations should be clearly defined. The business/functional requirement documentation should include both technology driven and non-technology 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. CSFs are those items considered to be either a business or technical requirement for the organization. The CSFs should enable the organization and the vendor to identify those areas of critical importance related to the successful implementation of the desired technologies.

Common examples of critical success factors from both a business and technical perspective typically include business-related and technical goals.

6.6.2 Business-related goals

Business-related goals include the following.

— Improved service: Users need the ability to quickly access and review information managed by the document imaging and workflow system.

— Compliance with regulatory requirements when applicable.

— Ability to track and monitor work activities: The system should enable 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 levelling 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, limiting 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.

— Satisfy organizational and/or government regulations pertaining to document retention: The use of electronic data storage should adhere to any laws and/or regulations covering the storage, retention and retrieval of information on electronic storage media.

— Decreased storage costs: The solution should provide the ability to use optical storage technology to reduce the overall cost of storing and retrieval of all “hard copy” information.

— Decreased costs for manual document management: The cost for manual document management should be reduced together 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 should 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.

6.6.3 Technical goals

Technical goals include the following.

— Scalability: The system should be fully scalable, allowing for an increase of the number of users and volumes of data without replacing primary system components. This scalability should be in the areas of increased memory, disk storage, optical storage, CPU speed and size, etc.

— Migration path: A clearly defined migration path should be fully supported by the proposed solution. This migration path should 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 should 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 should 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 should 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 should be commonly available throughout the document imaging and workflow industries, be fully supported by the selected product supplier, and have full user and/or development documentation and libraries.

— System maintenance. The system should be maintainable by the organization with support from the vendor/solution provider. The ability for the organization to maintain aspects of the selected technologies should be considered, together with the level of resource requirements and availability.

6.7 Technology evaluation guidelines

When evaluating appropriate technologies required or necessary to meet business and technical goals, the organization should consider several factors associated with the technology. The evaluation of the appropriate technology should include the following.

a) COLD/ERM: When evaluating COLD/ERM technologies, the organization should review the downloading, indexing and storage processing requirements. Additionally, the organization should consider the complexity of configuring the system to support new and/or modified report formats and indexing requirements. The ability of the technology to support simplified user access to data via a "query" screen and the ability to "cut and paste" information from a retrieved report or page to a standard office application should be considered. When evaluating COLD/ERM technologies, the organization should ensure that the system is capable of loading and indexing the daily work volume without impacting the users. This functionality of "loading" should include automated indexing based on templates defined by authorized users.

b) Document imaging: When evaluating these technologies, it is helpful for the organization to perform "site visits" to other organizations similar in size and processing, who have implemented the solution being considered. The purpose of these "site visits" is to gather information related to issues/problems encountered by other users that potentially have not been identified or addressed by the organization. During these "site visits", all aspects of document scanning, indexing and verification should be discussed. Overall system performance should be reviewed, together with ease of use and processing accuracy and organizational satisfaction with the product/solution and the product/solution provider.

c) Document services: These services enable users to manage electronic information independently of the tool used to create the information (i.e. word processing, spreadsheets, facsimile documents, etc.). Document services typically enable users to check documents "in" and "out" of information repositories; support document version control; and support document, group, and file level security rules. When evaluating these technologies, the organization should consider whether the product supports these functions together with being integrated with web publishing components (described below).

d) Workflow: When the organization determines that workflow technologies are required, it should be decided whether ad-hoc, administrative or production level technologies are required. For ad-hoc and administrative routing/workflow requirements, the organization should evaluate whether the product includes simplified authoring tools (for non-complex routing procedures) which can be used in a graphical environment together with monitoring capabilities. The monitoring capabilities should enable authorized users access to work queues or "baskets". These administrative and monitoring tools should further enable the authorized user to re-route work items and establish basic escalation and "time-out" procedures. These escalation and "time-out" procedures enable the users to establish a specific amount of time for which a work item can remain at any specific activity, or establish a total amount of time to elapse prior to automatically sending the work to a specific person or role. When the organization determines that production level workflow technologies are required, the escalation and "time-out" requirements should be included, but additional functionality should be considered. This additional functionality should include the ability for authorized users to build complex workflow rules and support load-levelling functionality and real-time work queue or "basket" monitoring.

e) Automated data capture: In many situations, the inclusion of OCR/ICR technologies can be justified solely on the reduction of manual data entry costs associated with indexing and capture of specific content from scanned documents. As there are many data capture products available that can be integrated with most document management systems, the organization should pay particular attention to the expected benefits and the ability to measure these benefits during the evaluation. When evaluating OCR/ICR/barcoding technologies, the identification of the following information may assist the organization in determining the expected cost benefits in comparison to manual data entry:

   1) colour of original documents and variety of documents or form types to be identified automatically,

   2) volume of hand-printed and machine-printed information to capture,

   3) volume of fields per form or document,

4) volume of characters per field,

5) field type (numeric, alpha, alphanumeric),

6) extent of document preparation (pre-sorted documents, mixed form types),

7) extent of forms re-design (dropout, bar code, OMR),

8) identification of business rules to validate or enhance the recognition result.

f) Forms management: When the organization determines that forms processing and management are required, the organization should consider both the forms creation and forms processing tools. The forms creation tools should enable the authorized user to develop new forms and modify existing forms for use within a browser-based application. This forms design should include the ability to create fill-in boxes, checklists, pull-down selections, free-form text input and digital signature attachment to the form during transmission. The forms management technologies should also enable the users to manage forms using version control and support the ability to either store the submitted data with the form or store the data with the version number of the form. This information should be stored in the application database for further management and/or storage.

g) Web publishing components: When the organization requires publishing documents to a web server, the system should support the ability for authorized users to create templates associated with specific classes or types of documents. These templates should be used by the web publishing system to convert submitted documents to either HTML or XML format including graphic and table conversion as required. The system should provide a mechanism for authorized users to either configure the system to automatically publish these converted documents directly to the web server or send the converted document to a webmaster for review and website updating.

6.8 Forms review and design considerations

Most organizations utilize both electronic and hard copy forms. These forms should be reviewed to identify where hard copy forms could be converted to an electronic format, and electronic forms should be reviewed to identify components that can be automated together with identifying components that should be stored/retrieved from legacy and EDM systems.

When reviewing electronic and hard copy forms, the organization should consider the following.

a) Does the form need to be an exact replica of the original hard copy version?

b) Does the form need to be printed and completed manually by outside personnel and then re-entered into the system?

c) Can the form be pre-printed with information for use by outside personnel (i.e. name, address, other fields completed from information already on file, etc.)

d) Can the form include bar codes or other “coded” information that would support automated indexing after return of the completed document by outside personnel?

e) Can the fields be streamlined and have components removed (i.e. received by, route to, date received, etc.)?

f) Can the form be provided electronically or does it need to be in hard copy format?

As organizations review forms, additional considerations such as whether to use forms creation applications designed primarily for web-based applications or forms creation applications designed primarily for EDMS environments. The use of electronic forms can greatly streamline organizational effectiveness and reduce duplicative data entry otherwise seen when using hard copy forms and/or forms that are completed manually and then scanned/indexed.

Organizations should consider the process that will be required to create the form, deliver the form to the users, and how to accept the form and/or information after being completed by the user(s).

6.9 Legacy data/document conversion methodology considerations

6.9.1 General

Legacy data is commonly loaded into the new EDMS systems including documents to be converted into a digital format and digitally born data. Issues associated with digitally born data loading or related to what versions should be loaded are how to establish common and accurate metadata. It is important to recognize that not all digitally born data will be stored in the system or stored in the system in the native format, due to age and/or format/structure of the data. For example, documents stored on a network drive or mainframe storage system may be in old or proprietary formats that require specialized formats/software no longer available or no longer easily accessible. These types of digital data should be converted to an industry standard format such as PDF/A, HTML or XML (or JPG, if image/map data).

For documents in hard copy format, there are three different approaches to existing file/data conversion in use throughout the document management/workflow industries: full back file, partial back file and as-needed. The organization should review and determine which approach best meets the previously defined business and technical goals. The approach selected by the organization will become extremely important if there are existing documents/files that need to be converted together with new and ongoing document receipt. Full back file and partial back file conversions typically require the selection of an outside "conversion" organization capable of processing large volumes of documents within a short time frame. The determination of whether to use an outside conversion organization or to convert using internal resources should be based on the volume of information to be scanned, the complexity of the required indexing and the required expediency of the conversion. The various approaches that should be considered by the organization include those in 6.9.2 to 6.9.4.

6.9.2 Full back file conversion

When selecting a full back file conversion, the organizational goal would be to have all existing hard copy documents available for use within the system in an electronic format. This conversion methodology is used when existing documents should be converted to meet business and/or technical goals. This methodology is typically very expensive and time consuming. The costs associated with full back file conversions are based on the volume of documents being converted and the total number of "keystrokes" needed to index each document, which is calculated by the total number of characters. When calculating the total number of characters, the organization should determine the level of accuracy required. For conversions where the conversion organization will only enter the information once (minimal data verification), the accuracy is typically not high enough to directly import the information into the document-imaging part of the system. It is recommended that a verification process (commonly achieved through "double-keying") be implemented, which increases the cost of conversion from an industry average of $0,10 per page to $0,20 per page.

6.9.3 Partial back file conversion

This conversion methodology is similar to the full back file conversion except that the organization selects specific documents requiring conversion such as by document age or date. Other than reducing the total number of documents requiring conversion, all considerations outlined within the full back file methodology apply.

6.9.4 As-needed conversion

This conversion methodology would allow the organization to convert documents only when required to complete an activity or process when new work is initiated. This conversion effort typically does not require the utilization of an outsourcing organization. To perform this type of conversion, the system should have a common "list" of where all documents are located, including both hard copy and electronic copies. The purpose of this list is to enable the users to quickly locate documents and determine whether they are available in the document imaging system or whether they are in hard copy format and require conversion. As new work items are received, the system should notify the user (or scan/index operator) that other documents

are in hard copy format and need to be retrieved, scanned and indexed prior to routing to the user(s) for processing.

6.10 Procurement document preparation

The procurement document provides detailed information on all aspects of the project. It is critical to include detailed technical requirements in any procurement document, including information on existing and anticipated operations, together with documenting data volumes, indexing requirements and routing requirements, etc. The technical requirements should include all necessary information allowing solution/product suppliers/integrators to respond in sufficient detail for the organization to be able to select the best product/supplier meeting the enterprise goals of the organization.

This document should also include detailed information related to the acceptance testing criteria used to validate components and solutions implemented by the selected vendors/suppliers. At a minimum, the software procurement documentation should include information on the following topics:

— business requirements;

— technical requirements;

— software vendor experience requirements;

— identification of standards with which vendor products should comply;

— description of how the organization will conduct acceptance testing;

— requirements of user acceptance testing documentation to be prepared by the selected solution vendor or product supplier;

— software vendor project management requirements;

— description of records management capability requirements;

— staff orientation and training requirements;

— requirements for vendor supplied technical support.

6.11 Solution/product evaluation guidelines

When evaluating solutions/products, the organization should consider several factors associated with the product and technology. Areas that should be considered include the following.

— Product maturity: The organization should evaluate the level of product maturity. This evaluation should include determining how long the product has been generally available, whether the product is in an early release stage (is this a new version which has not yet been fully implemented by the user community?), or whether the selected product has been in production for at least one year. All products are continually being updated to provide new functionality, “bug” fixes, and adherence to new standards and technologies. It is important for the organization to consider the maturity of each part of the selected solution when determining the overall risk factors associated with implementing these technologies.

— System configurability: The level of configuration capabilities of the system as compared to custom development should be considered. Most systems provide significant configuration capabilities, which greatly simplifies system design and implementation. This level of configuration is common for document imaging and document/library services and forms processing. Workflow technologies almost always require custom development to address legacy application integration. Other areas may also require custom development to incorporate functionality not readily available through a configuration process. Organizations should consider the level of configuration and custom development required to meet business and technical objectives and goals without adversely impacting user operations.