TECHNICAL REPORT

ISO/TR 22957

Second edition
2018-12

Document management — Analysis, selection and implementation of of enterprise content management (ECM) systems

Reference number
ISO/TR 22957:2018(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 171, *Document management applications*, Subcommittee SC 2, *Document file formats, EDMS systems and authenticity of information*.

This second edition cancels and replaces the first edition (ISO/TR 22957:2009), which has been technically revised. The main changes compared with the previous edition are as follows:

— updates have been made to bring the document in line with current generation technologies;

— the references have been revised and updated throughout;

— terms and definitions have been added and the acronyms section has been removed;

— "electronic document management system (EDMS)" has been changed to "enterprise content management (ECM)" throughout;

— the wording has been improved throughout and the contents have been reorganized to provide clarifications.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document provides detailed information associated with the analysis, selection and implementation procedures associated with enterprise content management (ECM) systems. The development of this document is a result of organizational requests to receive vendor-neutral industry information associated with technology standards, technical reports and industry best practices for ECM projects.

Terms and acronyms associated with various aspects of ECM 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, typically by incorporating additional levels of functionality and very rarely resulting in an entirely new core technology. This is important to note as the core ECM technologies are constantly maturing, and solution providers are identifying not only new approaches to addressing organizational issues and requirements, but also expanding the use of these technologies into areas previously unconsidered. 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.

For the purposes of this document, the terms "document management" and "content management" can be considered to be synonymous. As the ECM industry (previously referred to as the "document management" or "electronic content 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 (ESI)".

Clause 4 provides detailed information describing each of these technologies, and how they operate and inter-operate.

Clause 5 provides detailed information associated with currently available industry standards and technical reports.

Clause 7 provides detailed information related to industry best practices associated with all the customary project phases for ECM 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 licences, etc. This is very important, especially with most organizations carefully examining all expenditures related to all aspects of technology procurements.

This document provides detailed guidance to organizations considering the use of any of those technologies that comprise ECM [document imaging, document/library services, routing/business process management (BPM)/workflow, records management applications (RMAs), forms management, enterprise report management (ERM), etc.]. It should be noted and acknowledged that a complete records management programme set up against ISO 15489-1 is critical to any organization and is integral to any complete and thorough management plan associated with electronic information regardless of whether it is referred to as a "document", "record", "audio", "video", etc., internally by the organization.

Document management — Analysis, selection and implementation of of enterprise content management (ECM) systems

1 Scope

This document gives guidelines for a set of procedures and activities to be considered and/or performed by organizations when planning, designing and implementing various enterprise content management (ECM) technologies. The aspects or project phases range from initial business analysis through to vendor/integrator selection and technology implementation. The implementation of processes to manage electronically stored information (ESI) requires significant participation from the affected business units, if the content is stored and managed when created/received and controlled through the information life cycle following organizational policies and/or records retention and control policies are applied. As these efforts require multiple people with different disciplines, including technical teams, records managers and organizational management, this document has been prepared taking those perspectives into account.

This document is applicable to both in-house and outsourced systems, including cloud solutions. It can also be useful when dealing with specialized business systems. The term "enterprise content management (ECM)" (or "document management") used throughout this document is intended as an all-encompassing term referring to capture 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 (WORM) technologies]. This document provides information to users related to the technical reports, guidelines and standards that have been developed for technologies commonly available in ECM systems.

This document 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 document, those technologies can be reviewed if determined to be appropriate by the end-user organization.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 12651-1, *Electronic document management — Vocabulary — Part 1: Electronic document imaging*

ISO 12651-2, *Electronic document management — Vocabulary — Part 2: Workflow management*

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

ISO/TR 15801, *Document management — Electronically stored information — Recommendations for trustworthiness and reliability*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 12651-1, ISO 12651-2, ISO 15489-1 and ISO/TR 15801 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
library services
administrative components of the ECM system that handle access to information

3.2
expungement
process of removing a document from a system and leaving no evidence of the document ever having appeared on the system

3.3
semi-structured document
document that contains both some level of structure and organization along with completely unformatted text or text without any structure

EXAMPLE Letters, emails, forms with free-form text components, forms.

3.4
structured document
document that follows a strict structure or format

EXAMPLE Table, database record.

3.5
unstructured document
document that has no pre-defined structure or format and contains free-form text, which may or may not be similar to other documents of the same type, and varies greatly in structure, content, terminology and format

3.6
intelligent document recognition
IDR
technology that incorporates various methods of capturing and extracting information (or data) used to identify a specific type of document and data extraction with minimal (or no) user intervention required

Note 1 to entry: The most basic or traditional method of IDR has been in use since the late 1990s and incorporates the processing of barcodes, patch codes and other manual indexing methodologies. The current generations of this technology can also incorporate more advanced technology, techniques and algorithms to self-teach, i.e. to update the processing rules to classify or to extract data without user intervention, while others require linguistic algorithms, referred to natural language processing (NLP), to process unstructured content. IDR is a common marketing term used to describe the process of capturing content (scanned documents and digital born) and extracting content from the document within the limits of the technology being used.

3.7
natural language processing
NLP
technology used to determine and identify key words and phrases within processing audio data (e.g. call centres) and free-form text (e.g. the body of an email)

Note 1 to entry: This technology is able to reduce words to their base constructs and perform other actions, such as stemming, along with locating similar words or phrases without user intervention. This technology also varies greatly from standard IDR technology due to the ability to automatically update rules as determined by the users without the need for technical intervention. This technology is best suited for unstructured documents.

4 ECM technology

4.1 General

Many organizations still function in a hybrid environment of both electronic and hard copy environments. This environment is a direct result of the need of the organization to maintain information on all aspects of their business activities, as a business asset, to enable business efficiency and accountability. These requirements for organizational record keeping are often driven by government regulations and legislation to demonstrate accountability and implement quality control procedures. Organizations are also dealing with the exponential growth of digitally created content, ever increasing volume, diverging variety of formats, the acceleration of the velocity that content is arriving or being created and the growing doubt about the veracity of this content. These pressures lead to the disorganization of information as organizations try to keep up with the scale of change.

It is therefore a very important consideration for an organization evaluating or considering ECM technologies to first design, plan and implement the necessary foundational components, supporting the informational and organizational needs, and then expand and/or add additional functionality as required by the business units. This approach to phasing enables organizations to fully adopt new technologies while minimizing typically encountered issues adversely impacting day-to-day operations.

Organizations should also recognize that there is a wide range of stakeholders to be consulted, including senior managers, end users from business units, records managers and archivists, legal counsel and information technology (IT) staff.

4.2 Functional view of ECM systems

To help frame the concepts within this document, common terms used to reference these technologies should be discussed.

"ECM systems" 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);

NOTE Most ECM systems allow users to use this technology to also manage scanned documents, if desired.

— BPM/workflow (used to automate work processes, including the creation, routing, tracking and management of information being processed);

— ERM (used to store electronic formatted reports);

— forms processing (used to incorporate interactive forms and manage related forms data);

— OCR/intelligent character recognition (ICR) technologies;

— enterprise or faceted search;

— records management modules;

NOTE In some instances, this will be a module and in others an ECM will act as an integrated "front end" for an application that can be separate, such as RMAs, case management applications, legal discovery tools or web content management (WCM).

— metadata management;

— various applications (also considered add-ons), such as legacy system integration tools.

These systems provide users with greater access to ESI from common user interfaces, typically utilizing industry standard internet browser technology. One of the primary reasons users prefer web-based

client tools is the distributed functionality and ability to maintain standard desktop configurations for other office and business-related applications.

Figure 1 is a functional view of how ECM technologies, including records management, workflow, forms processing, etc., typically integrate within an organization. This view shows multiple applications integrated with and operating in conjunction with the ECM solution resulting in an overall trustworthy and reliable ECM environment. As Figure 1 shows, while there are many solutions available, it is important to remember that the policy layer is equally as important as the presentation and content layers.

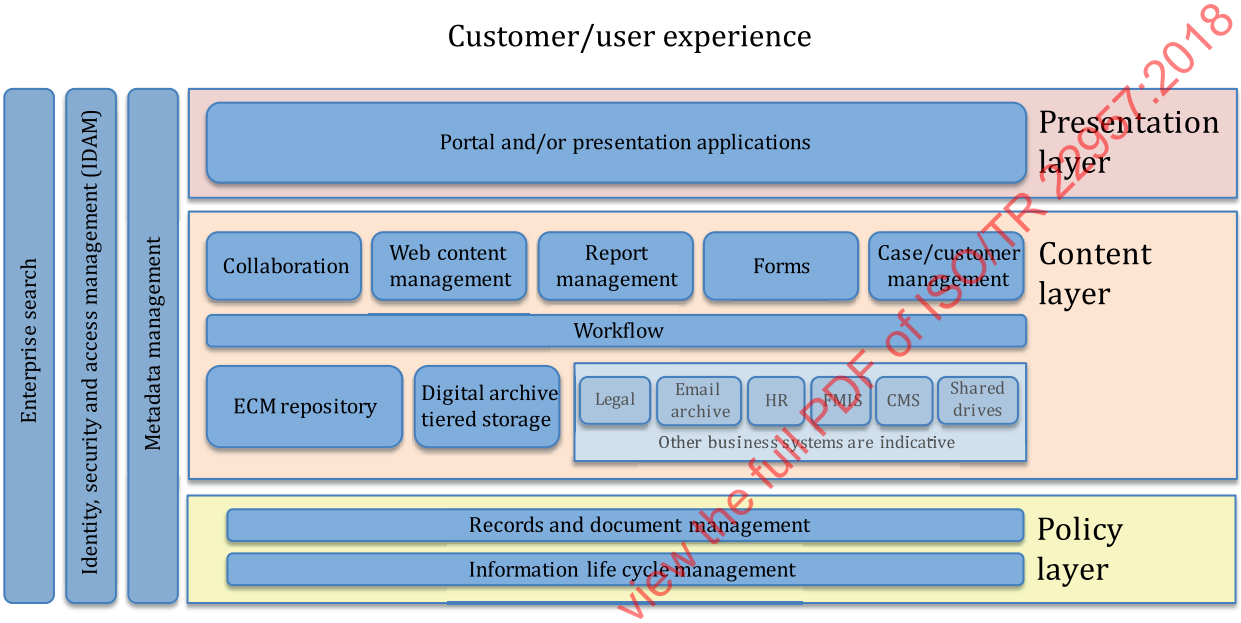

Figure 1 — ECM functional view

4.3 Technical view of ECM systems

4.3.1 Overview

From a technical perspective, the structure of ECM technologies can be viewed as a set of building blocks, as shown in Figure 2.

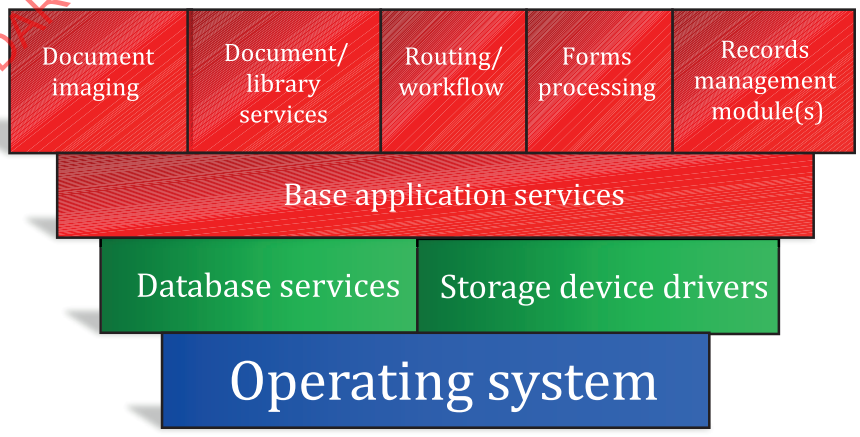

Figure 2 — ECM technology building blocks

The lowest level is the operating system. Database services and storage device drivers are installed onto the server as the second layer. The 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 (APIs), and application components integrating the core applications components with the database services along with providing the components that integrate the storage environment with the overall solution. The fourth layer incorporates the various core technologies of ECM technologies. Each of these core technologies (except for the operating system layer) are further described in the following sections.

4.3.2 Database services

There has been a significant shift from developing custom technology solutions at the database level to configuring/implementing commercially available software over the past 10 years. As the ECM industry and associated technologies matured, end-user organizations could shift from a "development" model to a "configuration" model for the base technological components. In fact, the selection of the database to be used is typically at the discretion of the organization but has become standardized using open database connectivity (ODBC) tools, which has resulted in the database components to be almost considered a "commodity item" rather than a specialized tool.

This is an important consideration for any organization evaluating ECM technologies from the perspective that almost all of today's ECM solutions have moved away from needing to have specialized database administrators to discouraging organizations from changing and/or modifying the ECM database table structures and configuration, which in many cases now results in the solution provider withdrawing solution support. While years ago it was important for the organization to hire dedicated ECM database administrators, this is no longer the case.

Since the very early 2000s, the ECM industry has effectively standardized 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 ECM solutions.

4.3.3 Storage device drivers

The storage device drivers or services are at the same level as the database services. The storage device drivers are used to connect the selected storage technology to the system and make the storage space available for the selected ECM solution.

4.3.4 ECM application services

The next layer in the "building block" is the ECM server application. Early in the development and maturity cycle of ECM technologies, end-user organizations were required to provide database administration and resources. During the late 1980s and early 1990s, ECM technologies had not matured to a level approaching commercial off-the-shelf (COTS) availability. While ECM technologies were maturing, end-user organizations were required to maintain the database along with the application.

Current versions of enterprise ECM solutions have shifted the database administrative functions back to the vendor/integrator with the end-user organization responsible for daily application maintenance and periodic server maintenance. Most enterprise ECM 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 using standard technology components configured to address specific environments and business needs.

The application services are provided by the selected solution provider and include all the necessary components supporting the platform of modules being implemented. These services include the various application interfaces to other applications, along with providing the integration points between the operating system security and management features and the management of the data being stored in

the ECM solution. The application services also provide all the core and underlying functionality for all the solution components.

4.4 Core technologies and application-specific modules

4.4.1 Overview

There are various core technologies and application-specific modules that provide specific functionality, including some or all the following core technologies: document imaging, document/library services, workflow, forms processing, records management modules, etc. All enterprise ECM 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/library services components into a single application, while other solutions simply integrate these components as required by the organization.

A very important aspect of ECM technologies is associated with trustworthiness and reliability. Any ECM solution should ensure that any information stored and managed will be retrieved exactly as stored without any unauthorized access, modification or deletion. Not implementing a solution that can protect the managed information is rarely of much use to an organization and becomes simply another data silo. Trustworthy ECM solutions incorporate not only software and hardware, but require organizational procedures and acceptance.

There are two very distinct approaches seen in the industry related to integrating and implementing various ECM technologies. There are applications that are what is referred to as "image enabled", meaning that the legacy application has a process to save scanned documents and records along with digitally born content (e.g. office documents, emails, facsimiles). Typically, the electronic (or digital) information is then stored on a network device or directly within the application/database (only for much older applications, typically). It is important to note that this is not considered to be an ECM solution.

On the other hand, there are ECM solutions that go beyond basic "image enabling" and provide full and extensive levels of control over electronically stored documents or records. These solutions always incorporate the following functionality, with varying degrees of sophistication and capabilities that allow for:

— capture of enterprise content (via document imaging or importing, if already in electronic format);

— process automation (workflows, forms processing, process automation, application integration);

— content management, storage and retention controls (trustworthiness, retention schedules, security controls, access controls).

The following sections provide detailed information on each aspect of these technologies and how they function and/or inter-operate within an ECM environment.

4.4.2 Document/library services technologies

Document/library services technologies enable organizations to manage digitally born documents and have become the core component of almost every ECM solution available in today's marketplace. A key component of document/library services is the ability to generate, log and track all aspects related to users accessing content, creating content and all associated actions taken within the system. Having this history generation and logging capability of all user actions when searching, accessing and/or creating content is the cornerstone of any trusted ECM solution.

These technologies control the authoring, check-in/out and/or version control of documents being developed, managed or stored. This enables collaborative development when desired, along with a mechanism to store/manage digitally born document libraries. The basic capabilities of these technologies include allowing authorized users to:

— load or import digitally born documents directly into the system;

— enter relevant metadata associated with the document, according to a metadata's schema previously defined in the organization;

— create virtual folders linking various documents, according to a classification schema previously defined;

— check information in/out of the repository;

— make changes and check the modified information back into the repository;

— manage whether original documents are updated or replaced during the update operations;

— establish security levels for groupings of documents.

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

Document/library services technologies ensure that any other user attempting to check the document out would, first, not be allowed to check it out, and, second, 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.

Along with the abovementioned authoring controls, document/library services are commonly configured to integrate with records management technologies, establishing retention controls and enforcing the records retention policies in an automated and regulated fashion, rather than allowing users to select relevant record series during indexing. The linkage between the retention schedule and the enterprise content is referred to as the "retention bridge", where information from the retention schedule is associated as appropriate with the classification schema.

NOTE The linkage between the retention schedule and the enterprise content is usually established by assigning retention rules to the elements of the classification schema (file plan).

Additional information on the integration of records management technologies is discussed below.

4.4.3 Document imaging technologies

Before discussing document imaging in general, it is important to note that terms used throughout the industry are constantly changing as the ability to automate information classification and data extraction continue to mature. Since the mid-1990s, the industry began exploring ways to streamline the imaging process and the term "intelligent document recognition (IDR)" became a common marketing term within the content/document management industry. Initially, the critical need was to incorporate software with the ability to deskew, despeckle, rotate and ignore certain type of pages based on colour and then these technologies matured/expanded to include barcode, patch code and other more advanced recognition capabilities. Throughout this timeframe, various solutions used varying terms to describe their solution and those marketing terms continue in various forms today, but the underlying process to reduce user intervention remained the same goal.

From a basic technology perspective, document imaging technologies enable users to scan hard copy documents into a digital or electronic format. The components of document imaging consist of a hardware device used to scan/convert the document into electronic format and the software used to control the scanning device and (in some cases) allow users to adjust the image quality to improve legibility (many of these functions are described above). Other solutions and/or applications simply integrated document imaging functionality into their core applications. It is important to note that this is not the same as integrating document imaging to an ECM solution.

When integrated to an ECM solution, the technologies enable users to index or enter metadata directly into the system, which then stores and manages the information, including historical activities associated with the document itself. Much of this information comes from an IDR process to some degree, which is described in more detail below. This metadata allows users to identify (or index) this digital information allowing them to be retrieved later and all types of information required by the end-user organization to fully track all necessary metadata. Some ECM solutions support more efficient indexing through the support of a controlled vocabulary that regulate synonymies and polysemies.

4.4.4 Intelligent document recognition

During the late 1980s and 1990s, the entire capture portion of the ECM industry was based on converting hard copy documents into electronic (or digital) format for further processing. To decrease the need for fully manual indexing, the industry began integrating the ability to process "separator pages" between different documents, and printing barcodes (also considered patch codes) to simplify separating the end of one document and the beginning of another, while being able to extract some basic information from either the cover page or the barcode itself, if located on the pages. During that timeframe, users (or "preppers") would insert a coloured page or page with a barcode in between each document, allowing the software to recognize the beginning and ending of pages and documents, thus beginning further automation of the categorization of documents being ingested.

Processing or extracting content from documents with some level of intelligence has been a desired goal within some aspects of the capture portion of the ECM industry, with the goal of reducing otherwise extensive user interaction to get the desired content into the library for later searching or delivery to a line of business application for processing. This process is commonly referred to as "intelligent document recognition (IDR)", which began as a marketing term to broadly describe data extraction from the technical perspective, relying on the appropriate underlying technology required, e.g. OCR for machine generated text, ICR for handwritten text, optical mark recognition (OMR) for marks, and natural language processing (NLP) or self-learning (i.e. when the solution updates rules internally based on user responses) for free-form text and audio, such as calling customer service. It is important to remember that each of these underlying technologies is entirely different and requires integration efforts from highly trained resources.

These technologies cannot be simply installed and turned on, as significant design and configuration work is necessary for the more advanced technologies, such as NLP and/or AI self-learning solutions, in order to achieve any reasonable level of accuracy. It is also important to note that different technology solutions incorporate different combinations of these underlying technologies to process structured or semi-structured documents (where the desired data can be identified in a consistent fashion). The use of these technologies is far more advanced than the traditional data extraction method used for structured documents or those that can be processed via manual data entry and/or using zonal OCR, barcodes, etc., which may be sufficient to meet the organizational needs. None of these traditional methods of extraction produce accurate results when processing free-form text, which requires the integration of linguistic and NLP (e.g. complaint forms, free-form contracts that contain some similar topics but in random locations).

Following the advancement in technology related to NLP, traditional methods were replaced with software based on linguistic technologies, primarily voice recognition and response, rather than extracting data from forms and documents. Linguistic and NLP technologies are also excellent tools to process free-form text (e.g. body of an email, a general complaint letter). The capture portion of the industry has continued to expand functionality to software tools that are readily available to all organizations, enabling users to automate as much as required while reducing user intervention, while other organizations attempt to fully automate the data extraction. The same technologies would not be a good fit for both examples as not all organizations have the need for AI "self-learning" or NLP. Traditional information capture with basic index automation is typically sufficient for many organizations. The latter technology still produces errors that must be managed by an operator, but, in time, the level of accuracy will increase and most capture software vendors will begin providing this functionality as part of their core products, just as almost all capture software vendors support

barcode recognition and zonal recognition where a user pre-defines where the software should find and extract the desired content.

NOTE Within the industry, the term "IDR" is used as an umbrella term to describe the various methods the system can extract content to classify and process documents being received. These technologies are almost always used with semi-structured and structured documents and rarely with unstructured information, as that relies heavily on NLP and the ability for the system to "learn" without user intervention to reconfigure and or create new templates for future processing.

4.4.5 Workflow technologies

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

When considering production workflow technologies, the organization should consider whether to employ role-based, user-based or both types of workflow technologies. User-based technologies require specific users to be assigned to specific tasks, while role-based technologies enable organizations to assign and re-assign users to groups or "roles" that are easily managed

Workflow provides for the automation of business processes and enables users to control the process logic, typically through a graphical user interface (GUI) and other tools. This ability to control the various business processes enables mission-critical, content-centric business applications to operate in an environment otherwise cumbersome to implement and manage. This has resulted in most ECM vendors offering an integrated workflow engine, or integrating the workflow engine with various workflow products readily available throughout the industry. The difference between these two approaches is whether the product consists of only those components developed by the primary product supplier or whether the primary product supplier has integrated specialized technologies developed by other suppliers.

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

4.4.6 Records management modules

Records management (RM) modules support and integrate into the ECM solution physical records management, electronic records management and/or litigation management features. Most ECM solution providers provide at least a basic level of records management capabilities while the larger, more established solutions include a wide range of records management functionality. Typically, the more advanced RM modules are used to provide litigation "hold" controls, management/legal override capabilities, the updating of retention periods and/or rules, and other features within the ECM solution. It is important to note that almost all ECM solutions have now fully integrated RM modules into their product suite and support many, but not necessarily all, functions required by records managers.

According to ISO 15489-1:2016, 3.14, records are "information created, received and maintained as evidence and as an asset by an organization or person, in pursuit of legal obligations or in the

transaction of business". Hence, records should be managed in a way that preserves their characteristics of authenticity, reliability, integrity and usability.

To ensure this, the ECM assigns and enforces the retention rules on documents and records either received or captured by other processes (such as receiving completed documents/records from another organization for historical purposes). This ensures that the overall solution manages the content securely, without undue duplication to comply with retention schedules and rules. Efficient configuration of these modules should be accompanied by a definition of responsibilities, records policy in the organization and tools, such as a classification scheme based on business activities, which has associated records disposition, security, an access classification scheme, a metadata schema and a controlled vocabulary, such as thesauri. This basis is established by senior management with the technical support of records managers and in cooperation with the CIO or organization's IT to implement an RMA.

Any RM module should have sufficient functionality for proper management and disposition of ESI throughout the information life cycle, and support capturing information in different formats. It should be noted that RM module functionality is a critical piece of an overall record and/or document management strategy for any organization.

In order to support records management policies and procedures of the organization, these RM modules provide configuration and management tools enabling organizations to implement rules to manage electronic records and control them automatically.

RM modules typically incorporate rules associated with a classification scheme based on business activities, records disposition authority and an organizational scheme of security and access related to the classification scheme. It is important that RM modules manage and control the preservation of the records identifying retention periods and subsequent disposition actions for specific types of information (regardless of whether referenced as a document, record, image, etc.).

There are many other functions and capabilities associated with records management, such as basic litigation management. Litigation management is a module that enables the organization to search, locate and maintain documents and records as required for legal purposes, and as required by regulatory or judicial authorities. When the records manager or legal counsel identifies documents/records that need to be maintained in adherence to legal proceedings, it is important to have the ability to prevent any users from accessing the information or simply to prevent users from modifying the information. This level of control over who should have access to the information should be at the discretion of the record management and/or legal counsel and typically based on the purpose for the litigation hold.

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

From a records management perspective, users should consult ISO 16175-2, which provides a good description of how content/records management technologies support records management disciplines and requirements

4.4.7 Enterprise report management technologies

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 disk, magnetic disk or magnetic tape as an alternative to paper printouts or computer output microfilm (COM). This formatted output consists of point-in-time reports, such as transaction listings of statements and invoices. Once this page output is stored on the ERM subsystem, it can be electronically retrieved, viewed, printed, faxed and distributed to workstations and host computer terminals within organizations or throughout an enterprise.

While there are many different data types in the computing environment, the type of data that ERM technology is concerned with is typically the result of transactions (data files and database records) being formatted by the application into a page-oriented form for printing on paper or COM. The structure and format of this output is known. These data are time-period focused; they are a snapshot of an internal system at a given point in time. These reports are often the basis for analysis or comparative reporting. 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, there are many complex tasks. Data should be downloaded or transferred to the ERM system server before they can be processed. The method used to transfer the data from the mainframe/host system to the ERM subsystem will vary depending on the communications capabilities currently in place. Recording consists of writing new documents to the storage media and then making them available for retrieval. Recording speeds vary from system to system and are most critical in high volume systems. The recording process involves:

— transferring the data to the storage subsystem from the host;

— processing the pages from the transferred file (i.e. extracting index keys, compressing and writing to electronic storage);

— 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.

4.4.8 Forms processing

The creation and utilization of electronic forms enable organizations to collect data in a standardized format and automatically enter or load the data into an ECM solution, with the data themselves being considered a record and part of the business process and important for the decision-making process.

Electronic forms are typically created using either a forms design package or using 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 usage of forms within the ECM industry has become widespread and most ECM 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. Using these tools, organizations are able to quickly develop and deploy forms driven data entry across the internet without significant development efforts.

4.4.9 Optical, mark and intelligent character recognition

Optical, mark and intelligent character recognition technologies enable organizations to quickly capture information from hard-copy documents (and some digital born documents that do not have renderable content, such as a document scanned into a format where the text has not been rendered) that needs to be captured during the indexing process. OCR, OMR and ICR can greatly reduce the time required to index documents while enabling organizations to develop in-depth full-text searchable databases. The differences between OCR, OMR and ICR are noted below:

— OCR: technology used to capture character by character text from documents;

— OMR: technology used to capture “marks” or “checks” from documents;

— ICR: technology used to capture hand-written information from documents.

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 ECM industry, it has been found that the use of these technologies can also greatly improve the quality of information being indexed, reduce the overall staffing requirements to perform the same functions manually, and significantly enhance use and adoption of the system.

5 ECM guidelines and standards

5.1 General

Industry guidelines and standards enable organizations to follow industry accepted practices and procedures. Standards and recommended practices specified in international, national, federal, state or local law or regulations are required specifically in the area covered by the law or regulation. In order to procure standard-compliant solutions, an organization should specify compliance requirements in its procurement documents and contracts, since this is the only way to make a voluntary (or legally required) standard obligatory for a vendor/integrator. Users of standards should also be careful to specify exactly what requirements in a standard need to 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. Likewise, it is possible that a standard or portions of the standard might not apply to the organization’s needs.

Following industry guidelines and standards will further improve the ability of an organization to implement the selected technologies according to industry’s best practices and standards. These guidelines and standards also enable organizations to implement products and technologies meeting their specific needs while being able to share information with other organizations that may use different solutions.

Industry guidelines and standards enable an organization to gain detailed information necessary to successfully prepare for, select and implement the desired technology. The guidelines that users should evaluate include:

— trustworthy content/records assessments, see ISO 18829;

— request for proposal (RFP);

— document preparation procedures for scanning/indexing;

— planning considerations for technology implementation;

— determining 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.

Industry standards include:

— document services integration and toolkits;

— workflow integration and toolkits;

— document imaging;

— optical storage.

Product suppliers should be required to 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 document will be updated to reflect those changes after the standards/guidelines have completed ISO approval processes.

5.2 Selecting the appropriate guideline or standard

5.2.1 General

It is recommended that organizations preparing to select document management and workflow products review relevant industry guidelines and determine whether the product vendors/integrators being evaluated meet the appropriate standards associated with that part of the technology. Titles of relevant standards and guidelines are provided in Annex A. Examples of various standards and guidelines for each of the technologies are provided below.

5.2.2 Industry guidelines

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

Prior to selecting a specific vendor/integrator, the organization should document relevant current and anticipated processes to be affected by the new technology, organizational and system requirements, and organizational expectations (see 7.3). This information should be provided to those vendor(s)/integrator(s) being considered. Regardless of whether the RFP is being sent to a single vendor/integrator or multiple vendors/integrators, the RFP should contain sufficient levels of information to enable vendors/integrators to clearly understand all business and technical goals and operational requirements.

This is very important as many (if not most) project failures are due to either misunderstandings between the organization and vendors related to what the final system should incorporate or a general lack of specific requirements that need to be met by the vendor/integrator. The RFP should contain sufficient information to allow respondents to fully understand the issue/problem being addressed, along with the technical and business requirements related to solution functionality. Providing details of the required functionality is important, as ensuring the inclusion of all the necessary components to establish a stable ECM technology foundation is critical to the overall success of the project.

5.2.3 Trusted system and legal considerations

5.2.3.1 Overview

Recognizing that all content management systems manage both record and non-record electronic documents 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 these data are called a “document”, “record” or some other term used by the organization, all ESI should be stored in a trusted environment when required and in compliance with the associated record retention schedule/plan.

Taking this into consideration and ensuring that all electronic information is stored and managed in a trustworthy and reliable fashion, compliance with the concepts given in ISO/TR 15801 and those related to the records management policies given in ISO 15489-1 should be considered. This will ensure that both technical planning, design and implementation along with records management policies and procedures result in the implementation and operation/management of a trustworthy and reliable ECM system for all ESI. 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 is maintainable in an ongoing fashion after rollout into production.

A trusted ECM system ensures that all ESI can be a true and accurate copy of the original information created or received regardless of the original format. The trusted ECM system should ensure that at least two (2) separate copies of the ESI be stored in a trustworthy storage environment that meets, at a minimum, all the following conditions:

a) the trusted ECM system should utilize both hardware and media storage methodologies that prevent unauthorized additions, modifications or deletions during the approved life cycle of the stored information, while storing at least one (1) copy in a safe and separate location from other copies of the ESI;

b) the trusted ECM system should be verifiable through independent audit processes that ensure there is no plausible way for ESI to be modified, altered or deleted during the approved information life cycle;

c) the trusted ECM system should maintain an unalterable audit trail that identifies the author and date of creation and any revisions to the original document.

It is important to note that trusted ECM systems incorporate not only hardware and/or firmware/software-based methodologies and technologies (software only security based solutions are not considered trustworthy by themselves), but also require adherence to organizational policies that ensure 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 ESI trustworthiness and reliability can be found in ISO/TR 15801.)

5.2.3.2 Expungement

It should be noted that information being expunged needs to follow specific legal rules and does not necessarily require that documents be permanently deleted, but can require that access to documents be permanently removed. Advice from legal counsel should be requested to determine whether making the information permanently inaccessible would be compliant with relevant legal requirements and remain compliant with ISO/TR 12037.

5.2.3.3 Legal admissibility of records

Local legal guidelines can apply for evidentiary issues associated with using electronic imaging systems and optical storage technologies, see ISO/TR 12654 for standard guidelines.

5.2.3.4 Retention requirements

Users and systems designers should consult the organization's established retention requirements given in their records management policies and procedures. The system being implemented should ensure that the system is able to retrieve the information throughout the required document life cycle. The storage media and their 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 and provide a mechanism to automate when documents are to be archived, for how long and what action to take after the retention period is passed, along with numerous other organizational advantages from a management perspective.

5.2.3.5 Redaction

The process of redaction may be elaborate, expensive and subject to judicial review (see ISO/IEC 27038). 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 a document viewer to discover the redacted content, and the supervisory review and approval of the redaction. A records management process is required 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 ECM system.

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

5.2.4 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 document.

5.2.5 Implementation considerations

5.2.5.1 General

Implementation considerations should include:

— system administration;

— security requirements;

— capacity planning;

— system performance;

— system scalability;

— fax services.

5.2.5.2 System administration

When selecting the technologies required 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 vendor 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 troubleshooting tools;

— hardware troubleshooting tools;

— database management utilities.

These technologies can be hosted off-site as well as on-site. The IT group within the organization should, if desired, be provided with the tools to perform these system administrative functions. At a minimum, the product vendor or system integrator should provide the ability for the IT group to manage the system and utilize the technical/support staff within the product vendor to resolve application and/or database issues that may be encountered along with assisting in software updates.

5.2.5.3 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 appropriate.

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 general operating system and other infrastructure security related requirements, organizations should review ISO/IEC 27001 and ISO/IEC 27002.

5.2.5.4 Capacity planning

To ensure an accurate assessment of the scope and size of the ECM system, the organization should determine the capacity requirements of the expected solution.

The following is a partial list of some sizing parameters to consider:

— system availability requirements;

— number of form types and documents;

— volume of existing forms, documents and records;

— 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.

5.2.5.5 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 system performance expectations be defined. The organization should determine the anticipated response times they expect from the system for:

— number of users accessing the system simultaneously;

— document retrieval from long-term storage media;

— document retrieval from online cache;

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

— document printing;

— scanning/indexing performance.

5.2.5.6 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;

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

5.2.5.7 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: Users can fax documents directly from their computer without viewing each document first. The user should have the ability to select a range of documents and have them routed to the fax server for transmission.

— Outgoing fax after document viewing: Users can fax a document during viewing. The user should have the ability to attach other documents to the outgoing fax as appropriate.

— Incoming fax processing: The system receives incoming documents and automatically routes them based on configurable rules (via a system administration interface), such as by the incoming telephone number, through forms or OCR processing.

— Fax status reporting: Users can view the status of and historical information related to faxes they have sent. This historical reporting should be based on user security rights, preventing users from accessing other users' history, while supporting users with higher levels of security to access all historical records.

5.3 Document imaging

5.3.1 User guidelines

5.3.1.1 Overview

User guidelines should be reviewed for document imaging technologies to assist the organization during all project phases from the planning phases through 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.

5.3.1.2 Planning

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

5.3.1.3 Indexing the information

As the organization plans to implement document imaging, the organization 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. While the importance of preplanning cannot be emphasized enough, organizations should be ready to adapt these guidelines and procedures as the system is implemented and new challenges and opportunities arise. As described in 4.4.4, advanced technologies can be utilized as determined appropriate, based on the types of documents being processed. IDR is a common term in use since the late 1990s to describe the inclusion of automating as much of the manual indexing as possible without introducing data extraction errors. It is also important to note that this technology is only of use with unstructured or semi-structured content as compared to the traditional method of data extraction using zonal OCR, barcodes and/or patch codes.

5.3.1.4 Storage technologies

During the planning stages of the project, the organization should review guidelines related to designing and managing a trusted ECM system (see ISO/TR 15801). There are various approaches to electronically storing information, including using optical WORM for long-term preservation to magnetic WORM for those organizations that need faster retrieval speeds and are less concerned about long-term archival issues to solid state storage technologies.

Users should exercise caution when using non-standardized or proprietary storage technologies. They should evaluate storage solutions against compliance with standards. Regardless of the storage technology selected by the organization, it is very important to consider the trustworthiness and reliability of the ECM system. This includes considering relevant government laws and regulations that require storage of the electronic information where it is not plausible that any stored information can be modified, altered or deleted during the information life cycle.

5.3.1.5 Formats

5.3.1.5.1 Overview

The organization should ensure that all information being scanned or electronically received is stored in an industry accepted format, such as JPEG, JPEG 2000, TIFF, PDF [see ISO 32000 (all parts)] or PDF-A [see ISO 19005 (all parts)]. Content may also be received in other electronic formats, such as audio (MP3, AAC) or video (MPEG) files. The MPEG file format has been standardized as ISO/IEC 13818 (all parts) and ISO/IEC 14496 (all parts) and should be used whenever possible. Use of non-standard or proprietary video, or audio formats should be discouraged unless there are no other alternatives. MP3 or AAC file formats, used for audio, have not yet been standardized. They are very commonly used and widely accepted.

5.3.1.5.2 Image formats

When raster image storage is the best choice for storing content, the organization should use either JPEG or JPEG 2000. Caution should be exercised when using TIFF files because of potential issues when rendering and reproducing these files in the future. When using the TIFF format, it is important to validate that the application generating and/or storing the TIFF files does not use non-standard headers or tags that are not well documented.

Image formats, e.g. JPEG, JBIG, JPEG 2000 and TIFF, can be modified with image editing tools without user knowledge. PDF, however, provides a means by which changes to a document can be determined.

5.3.1.5.3 Document formats

Unlike raster image format, PDF (and PDF/A, which is a subset of PDF) allows for the inclusion of not only raster image data but also vector and text content. When using PDF files, it is important to validate that all the additional information necessary to render the file (such as fonts) is embedded within the file. This is one of the key benefits of PDF/A. It requires that all necessary content, such as fonts and images, are embedded inside the file itself.

PDF and PDF/A also support a feature called incremental updates (see ISO 32000-1:2008, 7.5.6) that can be used to determine whether the content has been modified after the original authoring. For use cases (such as ECM) that require a greater form of validation, these formats also support the use of digital signatures and checksums.

5.3.1.5.4 Non-standard formats

Non-standard or proprietary file formats should not be used unless there are no other alternatives. Non-standard or proprietary formats include any formats used by a single vendor/integrator and not accepted as a standard file format at a national or international standard level. Proprietary file formats include, but are not limited to:

— file formatting that uses file-wrappers to encapsulate standard file formats within a non-standard structure;

— variants of the TIFF format that are not fully documented by the vendor/integrator and independently verified by the organization to ensure proprietary information is not contained in any of the headers;

— non-standard file formats not used by multiple vendors/integrators, etc.

Using PDF (or any of the PDF subsets) as the output format for any hard copy conversion to electronic format eliminates many of the commonly seen problems found with non-standardized formats. For example, when creating a PDF/A from an office application, the format incorporates all the necessary fonts and other information allowing for accurate renderings on systems, including those that do not include all the system fonts used when the information was created. Some of these benefits can be obtained with non-standardized and/or older formats, but usually require additional archival software and/or hardware to provide similar capabilities.

5.3.1.6 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 that the information is available and readable. The organization should establish a documented process to ensure that all documents are properly scanned and indexed, even if they are using fully automated data extraction, as it is important to have the ability to correct any established rules used to extract and process information if errors occur. This documentation should be followed by all personnel performing scanning and indexing, along with providing a mechanism for index data entry verification prior to document committal to the storage media and/or transmission to the business process.

5.3.1.7 Scanning quality control

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

5.3.2 Implementation considerations

5.3.2.1 Overview

Implementation considerations should include the following aspects of converting hard copy documents into digital format:

— document scanning;

— document scanning and indexing;

— scanning/indexing throughput;

— document image compression;

— post-scanning processing;

— OMR, OCR, barcode and ICR processing;

— quality control;

— query/retrieval display time;

— printing time.

Each of the above items is discussed in further detail below. Organizations should consider their current and future business needs/objectives when determining what aspects of document scanning needs to be implemented and/or incorporated.

5.3.2.2 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 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;

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

5.3.2.3 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 barcode 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 barcoding 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 time required to prepare the document for scanning;

— scanning 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 pre-set common fields (for indexing purposes) when scanning in batch mode;

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

5.3.2.4 Scanning/indexing throughput

The system should be capable of scanning either single- or double-sided documents using scanners capable of processing the daily work volume at the selected scanning site. This processing will include document preparation, scanning and indexing. The system should be capable of supporting low-, medium- and/or high-volume scanning capabilities depending on user requirements and the 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 that stations are available for use when needed.

5.3.2.5 Document image compression

Image compression/decompression should support ITU Group 4, LZW, 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 can require royalties and/or other fees to be paid on a periodic basis, which are usually based on the total number of pages converted into that specific compression format. These licensing/royalty issues do not occur with non-proprietary formats.

There are various compression methodologies available. ISO/TR 12033 provides guidance for users on selecting the appropriate compression technology that the vendor/integrator needs to support for different types of data. The different types of data may include scanned documents, line art, photographs, etc.

5.3.2.6 Post-scanning processing

Post processing may be used to provide image clean-up after 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.

The use of 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 can affect the ability to authenticate the document in a legal proceeding.

5.3.2.7 OMR, OCR, barcode 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 have produced remarkable benefits often evaluated in reduction of manual keystrokes. Evaluation criteria to consider when analysing the use of automated data capture include the following.

— 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? The amount of data to be captured and the cost to support a manual solution should be determined, and then it should be compared to an automated data capture solution.

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

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

— The business rules that may be used for post-recognition processing should be identified 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.

5.3.2.8 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;

— verify accurate security for each document.

5.3.2.9 Query/retrieval display time

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

The time periods include the total time required to retrieve the appropriate optical/removable media (when necessary), reading all requested pages from storage media, storage of all requested pages on a magnetic cache (if being used) and subsequent transmission of the first page to the user for viewing. When removable media (e.g. optical WORM, CD, DVD, tape) is implemented, the response time should consider the 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;

— retrieve information from the media.

5.3.2.10 Printing times

The imaging system should be capable of printing user-selected documents within anticipated user-established time frames. The 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.

5.4 Document/library services

5.4.1 Technology standards

5.4.1.1 Overview

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

5.4.1.2 Open source distribution

The product vendor/integrator should confirm 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.

5.4.1.3 Development toolkits

The vendor/integrator should confirm that the system uses industry standard APIs. 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.

5.4.2 Implementation considerations

5.4.2.1 Overview

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

5.4.2.2 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 should use a security model that ensures only authorized personnel can perform these functions.

5.4.2.3 Authoring controls

The system should provide the ability for the organization to establish authoring controls, including, but not limited to, allowing the author to determine what other types of users can access and/or modify the ESI. These controls should also enable organizations to establish collaborative development of ESI while ensuring a full and detailed history is maintained by the system. Common use of authoring controls is the creation and management of electronic forms and/or forms that are required to be used by various control or governmental agencies.

5.4.2.4 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 supports the ability for an authorized user to create a copy of a document within a specific folder, or set of folders, 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.

5.4.2.5 Group/user security

The system should provide the ability for organizations to apply security access/restrictions at both the 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.

5.4.2.6 Document security

The system should provide the ability for organizations to apply security at the document and/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.

5.4.2.7 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.

5.4.2.8 Document publishing to a website

The system should provide the ability for 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 use web templates to reformat the document into an HTML, XML or PDF-A format.

5.5 Workflow

5.5.1 Technology standards

5.5.1.1 Overview

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 grouped 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 provides 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 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. Several such interfaces were identified and became the foundation for the WfMC standardization program.

It was an important principle that the reference model focused 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 complementary.

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

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

5.5.1.3 Workflow auditing (WfMC interface specification 5)

The vendor/integrator should confirm that the product supports the WfMC audit specification (WFMC-TC-1015). This specification details information to be captured and managed by the workflow system during operation. This will ensure that all relevant data are associated with all functions within the workflow technology.

5.5.1.4 Workflow interoperability (WfMC interface specification 1)

The vendor/integrator should confirm that the product supports industry interoperability standards, including the usage of standard email systems. These interoperability standards will enable the organization to share workflow information directly between different workflow systems without requiring specialized development.

5.5.2 Implementation considerations

5.5.2.1 Routing

Workflow technologies include various types of routing, including ad hoc, administrative and production routing. Ad hoc routing enables the user to define a specific process for a document to follow for that document only. Administrative routing enables the user 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 user 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 processes the document accordingly, including the ability to support work timeouts, escalation and work reassignment.

5.5.2.2 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. Organizations should require a role-based system when implementing production workflow technologies.

5.5.2.3 Routing requirements

For those organizations requiring a 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 later;

— 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 specified date;

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

5.5.2.4 Graphical rule designer

The system should support the ability for authorized users to create and modify work rules associated with the workflow system. This ability should include graphical-based design and management tools that are used to create/modify work rules within a Windows[1)] or browser-based user environment.

5.5.2.5 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.

5.5.2.6 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.

5.5.2.7 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.

5.5.2.8 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 rule designer tool (see 5.5.2.4).

1) Windows is an example of a suitable product available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of this product.

5.6 Records management

5.6.1 Overview

The checklist in this subclause gives examples of more detailed system functions that a records management module might need to perform to satisfy each of the basic records management functions. There are many ways that each of these records management functions can be automated or implemented, so this checklist is meant only to be a starting-point for organizations to develop system specifications. Organizations should consider records management related functionality, such as that documented in ISO 16175-2. See ISO 15489-1 for additional information on records management functionality and concepts in conjunction with this document; it provides technical and detailed guidance associated with the selection, design and implementation of these technologies.

Users should note that these specifications are rather general, but they do provide a foundational set of capabilities to be considered to automate records management functions, establish records management control over electronic records, or otherwise improve existing records management systems.

Not all of these specifications will necessarily be implemented as automated system functions. Some functional requirements may be implemented through non-automated organizational policies, practices or records management procedures.

In this subclause, the term "user" refers to authorized users only. Different functions are permitted to different groups of users: administrative functions to records managers, retrieval functions to end-users, etc.

For all records systems, the records management module should:

— allow users to print and view all system management and control information: file plans, security assignments, disposition instructions, etc.;

— allow users to print and view the records themselves.

For all records, the records management module should:

— assign unique identifiers to records and their associated metadata;

— capture as much metadata automatically as possible and reliably link metadata to the records;

— capture metadata for records in a records management module, including creator, creating organization, author, recipients, subject matter, format, various dates (date created, date filed), a vital records indicator, etc.

5.6.2 Records capture

The relevant core records management components should enable the organization to support one or both of the following capabilities in compliance with the records management policy(ies).

— Allow import of records from other sources. This may involve format conversion for records that are imported from external information systems. In this case, records are physically captured and transported to a recordkeeping system.

— For those organizations that need to apply records management functionality to external systems, establishment of a link from the module to a record in those external information systems to establish records management control. In this case, physical transport of the records from one system to another is not required.

5.6.3 Maintenance/use

5.6.3.1 Organize records

For all records systems, the module should:

— allow implementation of an organization-specific scheme for how records are organized;

— allow users to create an organization-specific file plan and link the file plan to records categories or records series identified in retention schedules; include disposition instructions and specification of file cut off dates;

— Allow users to select record categories (records series) from the retention schedule and assign files to these categories;

— allow records to be linked to other records (e.g. a redacted record with its non-redacted counterpart, a final report with its earlier draft versions);

— allow users to create file folders, and to add, edit and delete categories assigned to these file folders;

— prevent deletion of non-empty folders from the file plan;

— allow users to add, edit and delete records retention schedules, and to freeze (i.e. withhold destruction) or execute records retention schedules;

— execute disposition instructions (e.g. move a group of records from active to inactive status);

— allow users to assign a status to records that prevents their destruction;

— allow users to specify in the system the organizational structure, organizational locations, and staff or unit to which records management responsibility may be assigned;

— import information from other sources (e.g. pre-existing file plans, box indices);

— allow users to specify identifiers for boxes, their contents, locations and related accession information.

NOTE For records management modules, which can also manage paper records.

5.6.3.2 Maintain records security

For all records management modules, the functionality of the relevant core ECM components should enable the organization to support the following capabilities in compliance with the records management policy(ies):

— manage the creation of document versions and revisions without losing original content unless specifically designed (such as the case of working drafts that are removed when the final document/record is created/approved and the drafts are of no further value); in the event drafts are to be removed, the retention schedule and policy should clearly describe this process;

— prevent any modification of a record's unique identifier once it is defined;

— manage deletion of indexes, categories and other pointers to records;

— calculate and maintain a checksum for records and their metadata or use some similar technological means of detecting any alteration to the record or metadata;

— provide audit trails of all add, update, delete, access and retrieval activity as configured by the solution and required by the organization;

— maintain appropriate backup copies of records and recordkeeping systems;

— provide adequate recovery/rollback procedures and rebuild procedures, so that records can be recovered or restored following a system malfunction.

5.6.3.3 Manage records access

For all records management modules, the functionality of the relevant core ECM components should enable the organization to support the following capabilities in compliance with the records management policy(ies):

— control access so that only an authorized individual is able to retrieve, view, print, copy or edit a record or other entities (e.g. metadata, file plan) in the record keeping system;

— permit the identification of individual users and groups of users and enable different access privileges to be assigned to individuals or groups; access privileges may limit access to selected records or groups of records and may limit access by selected individuals;

— maintain the integrity of redacted records and ensure that redacted material is not accessible.

5.6.3.4 Facilitate records retrieval

For all records management modules, the functionality of the relevant core ECM components should enable the organization to support the following capabilities in compliance with the records management policy(ies):

— allow searching on metadata, record content or assigned subject categories (using a controlled vocabulary);

— ensure that all access privileges (permissions and restrictions) are enforced on all retrievals;

— allow searching based on a combination of metadata, content and subject categories within a single query; query results that may be a list of records and their locations or may be the records themselves;

— allow retrieval of records and associated metadata and allow retrieval of records based on defined links (e.g. between versions of the same record);

— provide a sufficiently powerful range of search features and options as needed to meet various organization requirements; these can include wild-card or exact-match searching, proximity or adjacency searching, relevance ranking of search results, use of stop-words, limits on maximum size of results set from a search, query by image content, etc.

5.6.3.5 Preserve records

For all records management modules, the functionality of the relevant core ECM components should enable the organization to support the following capabilities in compliance with the records management policy(ies):

— ensure that all records can be read and accurately interpreted throughout their useful life in that system;

— enable migration of records to new storage media or formats to avoid loss due to media decay or technology obsolescence as described in ISO 13008;

— ensure that all captured metadata remains linked to appropriate records and is unchanged throughout the useful life of the records, including after migration to new media or technology;

— monitor storage capacity and utilization, and alert system operators when action is needed (e.g. to increase capacity, back-up system files).

5.6.3.6 Audit/oversight

For all records management modules, the functionality of the relevant core ECM components should enable the organization to support the following capabilities in compliance with the records management policy(ies):

— Create and maintain an audit trail (also called use-history metadata) for all records activity and system functions.

— Provide access to audit trail information in the fullest detail (e.g. each individual record access, including record identifier, time, date, and user).

— Provide summary reports of audit trail information (e.g. number of accesses).

— Track failed attempts of all records activity and system functions.

— Maintain audit trail information (e.g. number of accesses, details of individual record retrievals, attempts to delete a record) so that it can be managed as a record.

5.6.4 Final disposition of records

5.6.4.1 Destruction

For all records management modules, the functionality of the relevant core ECM components should enable the organization to support the following capabilities in compliance with the records management policy(ies):

— identify records eligible to be destroyed, based on records retention schedules and disposition instructions;

— delete records in a manner that they cannot be physically reconstructed or otherwise retrieved;

— enable a record to be kept of all record destructions, providing certifiable proof of destruction.

5.6.4.2 Transfer

For all records systems, the system should:

— identify records eligible to be transferred, based on records retention schedules and disposition instructions;

— export records and metadata (i.e. copy and subsequently remove them from the system), in a format acceptable for transfer to archives;

— enable a record to be kept of all record transfers, providing certifiable proof of transfer.

6 Considerations for hosted solutions — Trusted third-party repository requirements

6.1 General

For some ESI that is the result of business activities, it is essential to maintain its legal admissibility under certain regulations or international and national law. In order to keep the business intention intact, the solution managing the ESI should ensure the information maintains authenticity, integrity and reliability and is verifiable, as can be required by legal considerations, such as UNCITRAL. An organization should implement trusted systems and document procedures to control the creation, receipt, transmission, preservation and disposition of the ESI. However, it is complex and expensive for an organization to validate the authenticity of ESI that was created, transmitted, maintained or stored a long time ago.

One approach is to use a trusted third-party repository (TTPR) as described in ISO 17068. In this case, a client would submit an authentic copy of the ESI to the TTPR for storage and management enabling the certification of the ESI authenticity. This is because the trusted third party can provide qualified services to guarantee the authenticity of the digital records while maintaining neutrality toward all parties.

A TTPR is a set of services, systems and personnel that ensures that digital records entrusted to it by a client remain, and are asserted to be, reliable and authentic, with the aim of providing reliable access to managed digital records to its clients for the required retention period. The purpose of a TTPR is to guarantee the custody of the client's electronic information for the retention period, and to issue certificates for the purpose of legal admissibility.

A TTPR provides organizations the paperless benefits of trusted ECM related to document digitization and authorized custody. In addition, it can work as an intermediary to provide a secure exchange of digital records, ESI or electronic documents between creators, senders and receivers in many forms of electronic transactions, business transactions and operational workflows. It also can provide a public service by enabling secure electronic information exchange between individuals or organizations.

6.2 Trusted third-party repository services

In many jurisdictions, a client can secure the legal admissibility of ESI eliminating additional paper-keeping requirements through the use of TTPR services.

NOTE This ability is jurisdiction dependent and many other jurisdictions require hard copy information to be maintained.

A TTPR can provide an authorized custody service for a client's record or ESI and certify their TTPR authentication. To this end, a service contract between a TTPR and a client should clarify and check the state of the art of the client's ESI or digital records. The service quality and level of a TTPR should also be defined and checked.

A client's ESI or digital records may be created in the organization, received or transmitted from another organization or moved from other systems. After a TTPR acquires client's ESI or digital record, the TTPR should archive, display and dispose of it in a trustworthy and reliable manner.

The trustworthy and reliable manner of TTPR service should be examined and proven in a regular audit programme by interested parties (i.e. inspector, auditor, evaluator).

TTPR services are categorized into three major functions from the client's perspective:

a) trusted retention (archiving) and disposition:

   1) the TTPR should service the authorized custody of clients' ESI, electronic documents or digital records by applying a high level of security and storage technology;

   2) the TTPR should define the retention (archiving) period and time of disposition; the client and TTPR should contract the retention (archiving) schedule in the service level agreement;

   3) the TTPR should preserve the client's digital records, electronic documents or ESI in the form of a trusted archive information package (TAIP);

b) trusted submission and distribution:

   1) the TTPR should check the authenticity of clients' digital records, ESI or electronic documents and provide a client with a trusted submission information package (TSIP) module to achieve reliable submission to the TTPR as soon as possible after the time of creation;

   2) the TTPR should permit access and recall by the authorized user depending on legal considerations, such as information privacy, expungement or redaction;

3) TTPR services should distribute clients' digital records, ESI or electronic documents in the form of a trusted distribution information package (TDIP);

c) authenticity certification;

1) the TTPR should issue the authenticity certificate for the preserved copy that is verified as the client's authentic ESI, digital record or electronic document;

2) the TTPR does not have responsibility for the client's content;

3) the TTPR should identify the level of authenticity and issue the authenticity certificate.

6.3 Requirements of trusted third-party repositories

A TTPR should keep and protect the client's digital records, ESI and electronic documents in a secured manner away from any risks, and maintain and verify their authenticity, reliability and integrity for the requested service period of retention. For its quality management, the service, system and procedure of a TTPR should be required to be audited regularly by trained experts or auditors. The effective procedure for qualified TTPR service can be authorized and recognized mutually by a public-private partnership under international and national law.

For the establishment, a TTPR should require the following systems, equipment or facilities:

— digital record repository system;

— transmitter-receiver system;

— network system;

— time-stamping system;

— audit trail management system;

— network security system;

— access control equipment;

— disaster recovery facility;

— system for certificate issuance and validation of digital records;

— backup system.

For the secured custody, a TTPR should consider the service contract with clients, including the following items:

— service period;

— transmission procedure and method;

— types of repository service;

— security and data protection;

— information provision;

— compensation.

A TTPR should develop, maintain and actively monitor adherences to policies or procedures in the following areas:

— client management;

— administrator's role and authority management;

— network and security management;

— digital records management;

— operation of transmitted and received messages;

— audit record;

— data backup and recovery;

— security management;

— migration and receipt management;

— client system management.

7 Good practices associated with ECM project phases/activities

7.1 Assessing the existing records environment

The first step in any project is to assess the current situation related to records and/or document creation, receipt, management and/or disposition. ISO 18829 specifies what should be incorporated for a compliant ISO records/document technology and environment assessment to be performed. It establishes as methodology for trustworthy records assessment and identifies activities and operations that an organization needs to perform, or have performed, to evaluate whether the ESI is or was maintained in a reliable and trustworthy environment(s). These trustworthy records environments use content or records management technologies, commonly referred to as either ECM or electronic document records management (EDRM), and enforce organizational records management policies and schedules.

ISO/TR 15801 and ISO 15489-1 establish the standards and good practices associated with implementing trustworthy records/document management environments. However, a standard is necessary to define the methodology used to evaluate these types of records/document management environments regardless of what technologies are currently employed by the organization. ISO 18829 establishes the assessment methodology to be followed to identify the level of organizational compliance with these standards as related to trustworthiness and reliability of information stored in these environments.

ISO 18829 applies to existing or planned ECM systems and is recommended for most records or content management systems that have been in operation for any length of time, including the use of off-site and user-local specific storage or records silos. Establishing the existence of a trustworthy system is an important step in documenting the reliability of ESI maintained within that system or environment. ISO 18829 is designed for use by organizations evaluating the trustworthiness of existing record/document management environments. It identifies the mandatory activities and areas that need to be examined by internal and external professionals with a thorough technical and operational knowledge of the specific technologies and methodologies being examined, along with understanding record management processes and activities.

7.2 Change management

7.2.1 Overview

The key to success in implementing ECM lies in integrating the key elements described so far: technology, readiness, operations and culture. The best way to accomplish this integration is through the active participation and involvement of users and management.

7.2.2 Champion user participation

Participation from all levels of employees in the implementation process is an underlying theme of the following recommendations. At a minimum, key users (known as “champion users”) should be

identified throughout the organization. These champion users are typically senior or lead users who can provide input and feedback via a bi-directional communication model, enabling the ECM team and the users throughout the organization to be involved in all appropriate aspects of the analysis, design and implementation project phases. The benefits of employee participation are increased motivation, higher productivity and improved quality. Resistance to work changes can be lower when members of the group participate in those changes, especially when:

a) participation is a necessary but not sufficient means of reducing resistance;

b) participation is "a feeling of involvement on the part of people, not just the mechanical act of taking part in discussions".

Organizations that have left the champion users out of the planning, problem solving, analysis and redesign, or that have only marginally involved employees through random conversations and presentations, have been unable to tie together the four key elements: technology, readiness, operations and culture.

One of the best ways to ensure participation is through a champion user team selected by senior management. This team should include representative individuals from all levels and all key job functions as well as members of the technical project staff. The goal of this group is to jointly identify issues, problems and potential solutions related to work processes and procedures within the organization implementing these technologies. The formation of this team will alleviate many of the us-versus-them problems that arise when the technical staff, records management team or end-user representatives work separately from the other portions of the business unit in designing a system. This champion user team should begin its work with a one- or two-day training session to review the following topics:

— team charter, roles and responsibilities;

— project objectives and goals;

— change parameters established by senior management;

— methodology for work redesign that looks at both the social and technical aspects of work;

— problem-solving techniques, such as brainstorming, flow charting, cause-and-effect diagramming, etc.;

— development of effective teams;

— effective team behaviours;

— use of new technologies as support for the project.

The champion user team should continue to meet on a regular basis to accomplish the following tasks.

— User analysis: Identification of users, to determine the extent to which their needs are being met and to identify actions that can be taken to increase user satisfaction.

— Technical analysis: Documentation of how work flows, where problems occur and where these problems are first discovered. Also, determination of how imaging can impact the current workflows and which variances will be eliminated or reduced by imaging.

— Work redesign: According to information already gathered, a rethink of new workflow designs.

— System design: According to information already gathered, finalization of system specifications.

— Organization design: According to new workflows and system capabilities, determination of structural changes to be made in the organization, if any.

— Implementation plan: Identification of the steps and resources required to move from the current organization to the new image-enabled organization.

7.2.3 Change management programme

Once the data has been gathered and analysed, a targeted change management programme can be created. It is the follow-through on carrying out the change management programme that will be most critical to the success of the implementation process.

The change management programme should look for additional opportunities to involve champion users in all aspects of the ECM project. This can be done by establishing user involvement teams. These teams will develop strategies for managing the key issues that were identified through the organizational assessment. For example, separate teams might look at issues such as communications, training, and policies and procedures. Since many of the team members may not have had opportunities to help shape their organizations, the teams, such as the design team, will need some initial training in team building, problem solving and meeting management.

While each organization will develop a change management programme that is unique to its own situation, three broad areas that every organization should address are:

— communication (see 7.2.4);

— training;

— job design.

7.2.4 Communication

In most organizations, communication is not as good as it should be. While this is not usually detrimental to the business, lack of communication at critical times of change can lead to a failure of the change effort. The following suggestions can help with communicating during times of transition:

a) develop a communication strategy that includes the development of awareness and training material to fully describe the problem previously identified by end-users and management;

b) acknowledge input from users who were interviewed and who participated in the process baselining activity;

c) identify the following:

   1) what information is difficult to find?

   2) what challenges do the users encounter with the current process?

   3) who will benefit from technology-based change?

   4) what types of non-technology-based change is desired or required?

   5) what will change?

   6) what secondary changes will occur?

   7) what will be changing for the organization?

d) compensate for losses;

   EXAMPLE In ECM environments, many employees feel a loss of socialization with their peers or feel that the use of these technologies will increase management oversight and force users to account for their time. While this can be an objective, how this is presented to the end-users will greatly affect their interest and willingness to participate in any type of change.

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

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

Some successful ECM installations have relied on monthly or quarterly meetings, newsletters, electronic mail updates, website communication pages, hot lines and suggestion boxes. Each organization should determine what will work in its environment.

7.2.5 Project phases and activities

7.2.5.1 Part 1 — Strategic alignment — Senior management participation

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

The strategic planning session should focus on answering the following questions:

— What aspects of our business culture are effective?

— What aspects of our business culture are ineffective regarding ECM implementation?

— How will ECM impact our employees and key external organizations?

— What structural changes in the organization are likely to result from ECM?

— How much change do we want in this organization? When examining the continuum of control versus commitment, where are we now regarding management style? Where do we want to be?

— What technology-based change is appropriate for the organization?

— How should the project be phased to allow adequate time for change management and organizational acceptance of the selected technologies?

— Do we simply want to automate existing processes, or do we want to fundamentally change workflows?

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

7.2.5.2 Part 2 — Strategic project assessment and planning (see ISO 18829)

The organization should have a full and compliant assessment of the current state of record and document management within the organization.

The output of this standardized assessment is a detailed report providing sufficient information allowing the organization to determine how best to address any areas identified as not being in full compliance. The report should also include, with detailed technology (if appropriate), recommendations and records/document management related policies and procedures required to come into full compliance.

A key element of this assessment is to provide detailed information to the organization related to the overall trustworthiness of their ECM environment, along with recommendations on how to address those areas evaluated as not following the associated ECM and records management related standards.

Upon conclusion of any compliant ISO 18829 assessment, the assessment team should prepare a detailed report containing, at a minimum:

— a business needs and/or business case section, including a description of the records assessment process followed, a summary of findings for physical records and electronic records, and business-related issues identified throughout the assessment;

— an analysis section providing detailed information following the standard records management models established by ISO 15489-1 and information related to accountability, transparency, integrity, protection, compliance, availability, retention and disposition;

— a technology gap analysis section providing a description of all relevant ECM, records management, and other document/record related storage or creation technologies currently in use by the organization;

— section on technical and records-related recommendations, including recommendations associated with changing the existing state of records management to establish a trustworthy ECM environment in compliance with ISO/TR 15801, ISO 15489-1 and other relevant ISO standards and good practices.

7.2.5.3 Stage 1 — Performing the detailed assessment

7.2.5.3.1 Architectural/operational gap analysis

All systems that are vital to the success of the ECM project should be identified and the potential integration points should be mapped out. The functionality desired to reach the end state should be defined and it should be determined if it is functionality that is part of the software package or if it a customization. It should be determined how workflow, rules, analytics and compliance play into the project and departmental requirements. The subject matter experts (SME) role and how they perform their job on a daily basis to identify areas of process improvement and/or automation should be evaluated.

7.2.5.3.2 Defining business objectives from a management perspective

Business objectives should determine the ultimate goals of the organization.

EXAMPLE Service level improvements, increased productivity, quality of service, preparing the organization to prepare for an electronic environment.

7.2.5.3.3 Identifying organizational goals from a management perspective

The management team should provide input and help define the goals of the organization in relation to expectations for the ECM project.

EXAMPLE Increased profit, stakeholder value, global presence, service quality excellence, streamline processing.

7.2.5.3.4 Defining how project success or failure will be measured

The introduction of new technologies means more cost to the organization. In order to assess whether organization objectives are met, all project deliverables should have measurable criteria, which may include service level improvements, such as reduced complaints from clients, reduction in cost of operation, cost containment on system deployment or reduced time to complete processing.

7.2.5.3.5 Determining technical objectives and goals

Apart from the business objectives, the technical objectives and goals should be determined based on organization IT infrastructure and capability to launch the system, and the cost of new technologies.

7.2.5.3.6 Documenting business requirements and expectations

The business objectives used to determine whether the solution meets the anticipated results should be clearly defined.

EXAMPLE Improving the service, being able to track and monitor the activities of work, increasing the efficiency of resources, meeting organizational and/or government regulations related to document retention, lowering the cost of manual handling documents.

7.2.5.3.7 Defining technology requirements

The technical targets favourable to the organization should be identified.

EXAMPLE Scalability of the solution, migration route, modularity, access via web, use of industry-standard components.

7.2.5.4 Stage 2 — Technology identification and project scope definition

7.2.5.4.1 Evaluating and assessing current technology

Technologies available in the market for document handling that can contribute to the objectives and business requirements should be investigated.

EXAMPLE Document imaging, document services/library, workflow, OCR/ICR and full text search, COLD/ERM, automation of the data capture, automatic identification of documents, forms management, components of web publishing.

7.2.5.4.2 Identifying non-technology-based change

A common outcome of the process analysis is the identification of processes, activities or procedures that are no longer required by the organization. Typically, when these processes, activities or procedures are identified, the organization changes the process to eliminate or streamline where possible without the use of technology, as the problem was not technology driven.

7.2.5.4.3 Identifying technology-based change

Along with identifying areas of non-technology-based change, the organization should consider what portions of the process could be improved and/or streamlined using various components of ECM technology. This technology-based change commonly includes either document imaging or document/library services, and quite often includes workflow and forms processing along with web publishing.

7.2.5.4.4 Establishing a champion user team

The success of any system is dependent on the people who execute the project goals. This will require a champion or champion user team to be identified who will carry out the overall project objectives.

7.2.5.4.5 Developing a high-level process baseline (see 7.3)

The process baseline will likely change to accommodate new technologies. In order to take advantage of the new processes, a high-level process baseline should be developed from current processes to new processes.

7.2.5.4.6 Preparing a detailed process baseline

As the high-level process baseline is examined, the detail process baseline will emerge. The detailed process should be recorded for subsequent examination of the system design and should contain sufficient levels of detail related not only to tasks and processes, but also to how information moves between tasks and how information is handled, processed, tracked, logged, etc.

7.2.5.4.7 Preparing anticipated ("to be") process steps/activities

The organization should identify all tasks that support the scope of the project and show the scope in detail, while also help to measure the development progress of the project.

7.2.5.4.8 Documenting specific processing metrics

The organization should document metrics to help it to display the advantages of implementing a new technology versus the earlier situation without using the technology.

EXAMPLE Cases processed per unit of time, time required to locate a document, information requests served within a period of time.

7.2.5.5 Part 3 — Vendor/integrator selection/procurement

7.2.5.5.1 Developing a detailed request for information (RFI, RFP, RFO, etc.)

The organization should organize and prepare all the information collected in Section 1 and should add detailed technical requirements clearly defining what the vendor/integrator is expected to provide and what functions and features are required. The level of technical detail should take into account various functions and capabilities identified and discussed within this document.

7.2.5.5.2 Identifying potential vendors/integrators to be considered

The organization should evaluate various vendors/integrators to identify those solutions that have a proven track record within the specific line of business associated with the business unit, along with sufficient resources to support any selected products/solutions. While there are numerous vendors/integrators in the industry, the organization should consider those vendors/integrators that have a demonstrated track record in successful solution implementation and long-term support. Vendor/integrator financial stability and local team resources to the organization are also of significant importance, along with technical knowledge and the strength of the local or other resources assigned to the project.

7.2.5.5.3 Evaluating vendor/integrator responses (creating a shortlist)

The organization should evaluate the list of vendors/integrators identified in Part 2, Step 2 to create a short-list of vendors/integrators considered to be capable of addressing the previously identified project objectives, goals and technical requirements. Other considerations for the vendor/integrator shortlist is project methodology and ability to work with organization. This can be initially evaluated by speaking with client references within the vertical market that is the primary focus of the organization.

7.2.5.5.4 Requesting vendor/integrator demonstrations/briefings

Each vendor/integrator on the shortlist should provide a detailed demonstration of how their product/solution could be used to meet the organization's objectives and goals. The demonstrations should focus on how the solution is used by other similar organizations and how the solution met similar client objectives and goals, with a minor emphasis on vendor/integrator marketing and sales literature. Vendor/integrator financial stability and local team resources to the organization are also of significant importance.

7.2.5.5.5 Selecting a desired vendor/integrator

Including representatives from the user community and the IT team, the management team should select the vendor/integrator that will best meet the organization's objectives and goals, and that will be able to operate within the organization. A good match between the methods of operation of the two organizations is typically of value.

7.2.5.5.6 Establishing a vendor/integrator statement of work including acceptance criteria

The organization should prepare a complete agreement with the selected vendor/integrator that clearly defines expectations and how the deliverables will be reviewed and accepted.

7.2.5.6 Part 4 — Solution design

7.2.5.6.1 Establishing project team oversight and management procedures

The team of resources to oversee the project should include resources with experience of implementing these technologies and who are able to provide subject matter expertise to the organization. The team should include senior management, IT and records management representatives, and an end-user champion. Depending on the size and complexity of the organization, the team might also consider legal, risk management, privacy and security, and other information governance representatives. The oversight resource should work with the organizational management team to monitor the work of the vendor/integrator, ensuring the solution is being designed following industry good practices and standards and that it will meet the objectives and goals of the organization. The vendor/integrator should be notified early in the procurement process if a subject matter expert will be directly involved in providing guidance to the selected vendor/integrator.

7.2.5.6.2 Vendor/integrator prepares the system and detailed design documents

The vendor/integrator should prepare a full system design document describing all the functions of the system and how they will operate. After the system design is agreed upon, the vendor/integrator should prepare the detailed design documents that will be used by the developers and configuration/installation staff to build all aspects of the system.

7.2.5.6.3 Organization finalizes the detailed user acceptance criteria

The organization should develop a detailed user acceptance document based on the system design documentation and anticipated functionality, ensuring that all components requested in the RFI/RFP/RFO and those bid by the vendor/integrator have been developed and/or implemented as expected by the organization.

7.2.5.6.4 Vendor/integrator develops the prototype system

The vendor/integrator should develop the system in a non-production environment. As the solution nears completion, the vendor/integrator should establish the prototype system, enabling the client organization to fully evaluate the functions and features. This should be functionality testing and not testing to resolve programming/configuration issues. The vendor/integrator should ensure that full and detailed testing is completed prior to the client evaluating the prototype system's functionality.

7.2.5.6.5 Organization evaluates the prototype system

As the organization reviews the prototype system, the users should make a record of the areas meeting their expectations and those needing to be updated or changed. It is important to note that since many users will not have used the technologies before accessing the prototype system, it is common for users to require changes after reviewing the prototype.

7.2.5.6.6 Vendor/integrator updates and completes system development

The vendor/integrator should complete system development and ensure that all design documentation is updated. It should continue updating the test environment to enable the organization to continue a process of review and feedback. This will greatly eliminate re-work due to misunderstandings after the system has been fully developed.

7.2.5.6.7 Vendor/integrator fully tests all aspects of the system

It is very important for the vendor/integrator to fully test all aspects of the system, ensuring that the system functions as planned and as expected by the organization. This testing should verify that all configurable components have been configured as planned, along with full testing of all functions or capabilities that were developed. In addition to detailed unit testing, the vendor/integrator should perform detailed end-to-end system testing.

7.2.5.7 Part 5 — Solution planning and rollout

7.2.5.7.1 General

Each of the following steps are executed in a repetitive fashion throughout the rollout of the selected technologies. This is important to consider, especially if the organization desires to implement these technologies to store and manage information upon receipt or to use these technologies solely as an archival appliance. While both approaches provide value from different perspectives, the following steps represent those activities that the rollout or implementation team should consider and continue to monitor through project rollout and technology expansion.

7.2.5.7.2 Implementation planning and change management considerations

Through all aspects of the project, the organization should include representatives from the business unit, IT and the management team to plan and discuss all the specifics of how the technologies will be implemented and what type of process and organizational change is required to support the new approach to conducting business. This planning is critical to the success of the project and should include establishing stepping stones to implement the technologies in phases, thereby reducing potentially adverse impacts to ongoing or daily business activities that need to continue.

7.2.5.7.3 Technical and user training

Both technical staff and end-users should be fully trained in the use of the system. The technical training should be sufficient to enable the technical staff to support the day-to-day end-user activity with the support of vendor/integrator. The users should be trained in detail prior to final system testing and during the final implementation planning phases of the project. The closer the user training is to the actual implementation date, the better for the overall project experience.

7.2.5.7.4 Business practice documentation preparation

The organization should develop a detailed business process document that clearly documents how information is received, stored and managed, along with the processes, policies and procedures that will be followed. This business practice documentation should be completed when the system goes into production and should be updated as the processes change, including documenting whenever the processes or procedures change.

7.2.5.7.5 Organizational policy creation/updates

The organization should ensure that all organizational policies are current and accurately reflect how the users will use the system. These policies and procedures should be identified in the business practice documentation, along with a description of how the organization ensures all affected users are notified of the policy and/or procedure.

7.2.5.7.6 User acceptance testing

The organization should ensure that adequate time is set aside for full solution and full user acceptance testing. Full solution testing typically includes actual users loading content. As users begin this process, users and organizations will always identify areas of the solution that need to be modified and/or updated prior to further production rollout. Once this process continues, it is important for the vendor/integrator to recognize that the organization still has other ongoing activities, which usually results in

greater time being required for full user acceptance. When the users fully adopt the technology, there should be a user acceptance test that demonstrates that the functionality is delivered and is being utilized. Usually, there are components that are not fully rolled out during the production startup, which should be noted and completed as appropriate.

7.3 Process/procedure baselining

7.3.1 Overview

The purpose of process/procedure baselining is to clearly define existing processes/procedures and to 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:

— how documents and information are received;

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

— 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 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 baselining process continues, users may describe processes and/or procedures that are not officially sanctioned in the day-to-day processing. These workarounds, or alternative methods, should 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 because 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 ECM industry related to process baselining. These activities include the development of a high-level baseline to establish the overall structure of the business process, a detailed baseline to document the specifics of each task/procedure identified during the high-level baselining, and associated manual processing metrics.

7.3.2 High-level baseline

The first step in documenting business processes is to develop a high-level (management view) of the manual or document-based business processes. This high-level view of the major processes and activities to be examined is also referred to as a “task oriented baseline”. These high-level baselines provide information related to the tasks performed by the organization without the detailed information associated with how information enters the process, how it is managed and how the information moves to the next stage of processing.

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 that document the manual, or document-

based, processing flow through the organization. These high-level schematics are then further exploded during the detailed process documentation, which provides the full set of information necessary to identify all aspects of the process for analysis and evaluation. Detailed information on developing a process baseline is given in ISO 10244.

7.3.3 Detailed process baselining

Upon completion of the high-level processing schematics, areas of the high-level baseline are typically identified by managers and supervisors to be further documented during the detailed process baselining process. Users throughout the organization are further identified to be interviewed to collect sufficient information to fully document the process at a detailed level. 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 a computer. A read-only version of the software should be provided to any organizational resource who needs to review this documentation. Users identified during the initial project planning stages and others identified throughout the process should be interviewed to ensure that all the relevant portions of the business processes are documented. All aspects of how information flows and is tracked, logged, managed, routed, along with any other work-related activity associated with the process, should be documented.

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 detail gathered for each process, including rules and conditions, is shown in Figure 3.

7.3.4 Processing metrics

A processing metrics report should be developed, including information related to personnel time spent on all work-related manual activities identified through high-level and detailed schematic development. These tasks can then be further examined by the organization 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, along with the anticipated time after new/updated technology implementation. Time associated with processes/activities documented in the schematics should be identified and documented to identify items, such as:

— time spent logging receipt of documents;

— time spent copying, filing and 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 documenting it, while comparing the received information with overall times for processes and other aspects of organizational work-related activities.

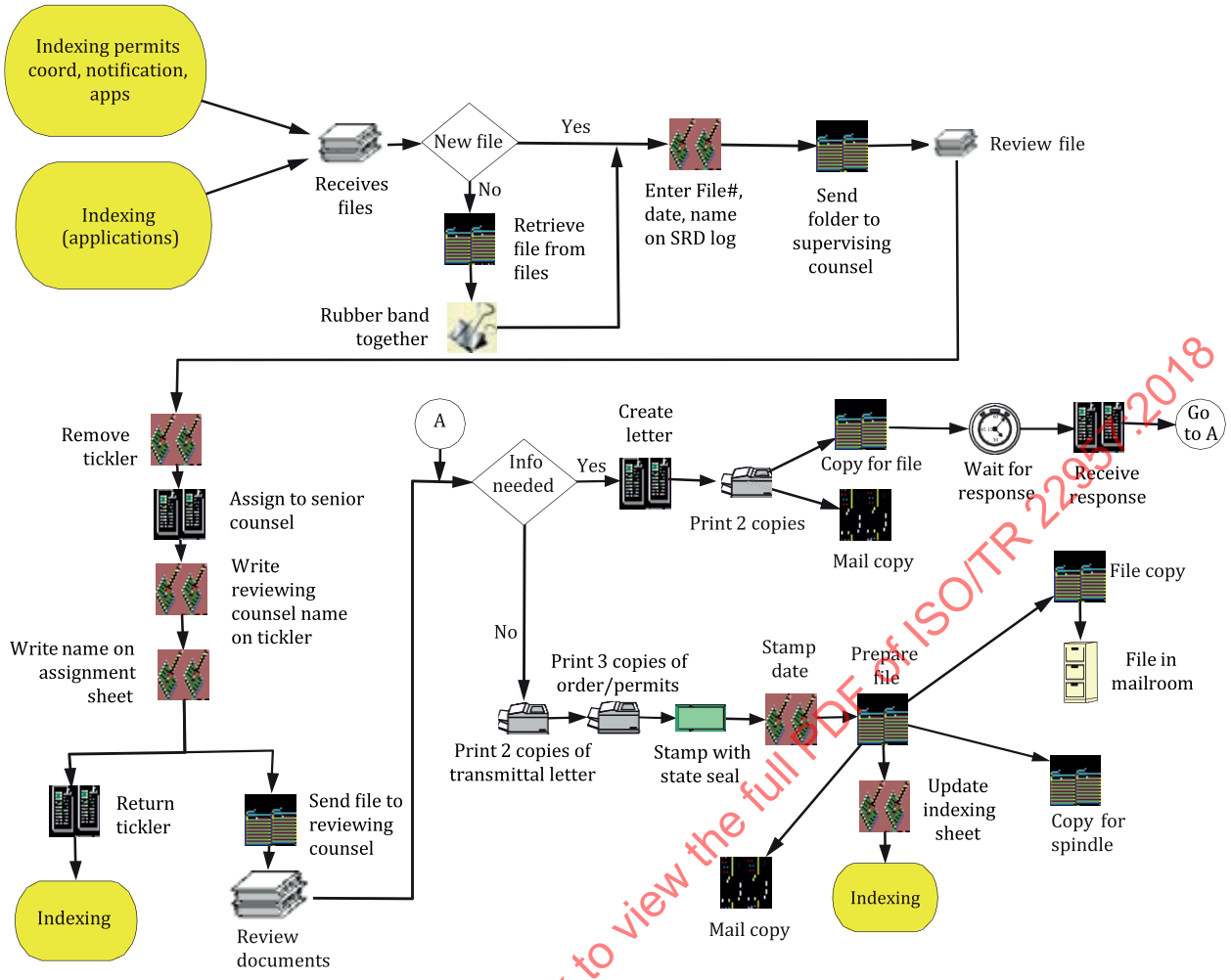

Figure 3 — Sample detailed process schematic

7.4 Anticipated processes/procedures

Upon completion of the baselining process, the 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 ECM 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.

7.5 Technology requirements definition

After identifying the relevant technology-based changes required by the organization, solution requirements should be documented, providing detailed information to potential vendors/integrators. This document should clearly define anticipated user and system functionality in sufficient detail to enable a potential vendor/integrator 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.

7.6 Document classification and indexing model

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

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

— fully addresses records retention issues and schedules;

— enable staff to easily locate and retrieve documents;

— support the ability to implement document security;

— provide logical information groupings;

— 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.

7.7 Business objectives and requirements

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

— the business objectives of the project;

— the business functional requirements;

— the 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 the business, technical and management team members of the organization.

Additionally, the definition of critical success factors (CSFs) 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. CSFs should enable the organization and the vendor/integrator to identify those areas of critical importance related to the successful implementation of the desired technologies.

7.8 Technology evaluation guidelines

When evaluating appropriate technologies required or deemed 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 factors.

— 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.

— 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, along with ease of use and processing accuracy and organizational satisfaction with the product/solution and the product/solution provider.

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

— Workflow: When the organization determines that workflow technologies are required, it should decide 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) that can be used in a graphical environment along with monitoring capabilities. The monitoring capabilities should enable authorized users access to work queues, often called “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 escalations and time-out procedures enable the users to establish a specific amount of time in which a work item can remain at any specific activity or to 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. Additional functionality should also 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.

— 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 ECM systems, the organization should pay 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:

  — colour of original documents and variety of documents or form types to be identified automatically;

  — volume of hand-printed and machine-printed information to capture;

- volume of fields per form or document;
- volume of characters per field;
- field type (numeric, alpha, alphanumeric);
- extent of document preparation (pre-sorted documents, mixed form types);
- extent of forms re-design (dropout, bar-code, OMR);
- identification of business rules to validate or enhance the recognition result.

— Forms processing: 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. These form design tools should include the ability to create fill-in boxes, checklists, pull-down selections, free-form text input and a 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.

— 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.

7.9 Forms review and design considerations

Most organizations use both hard copy and electronic forms. Hard copy forms should be reviewed to identify where they can be converted to an electronic format. Electronic forms should be reviewed to identify components that can be automated, along with identifying components that should store/retrieve from legacy and ECM systems (see ISO 12029).

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

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

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

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

— Can the form include barcodes or other coded information that supports automated indexing after the return of the completed document by outside personnel?

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

— 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 ECM environments, should be considered. The use of electronic forms can greatly streamline organizational effectiveness and reduce duplicative data entry related to 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 accept the form and/or information after being completed by the user(s).

7.10 Legacy data/document conversion methodology considerations

7.10.1 General

Legacy data are commonly loaded into the new ECM systems, including documents converted into a digital format and digitally born data. There will be issues associated with digitally born data loading or related to what versions should be loaded and 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 or software that are no longer available or are 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). Organizations should consider the methods and recommendations given in ISO 13008.

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 along 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 full back file conversion (see 7.10.2), partial back file conversion (see 7.10.3) and as-needed conversion (see 7.10.4).

7.10.2 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 need to 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 can significantly increase the cost of conversion.

7.10.3 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.

7.10.4 As-needed conversion

This conversion methodology allows 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

