TECHNICAL REPORT

ISO/TR 15801

First edition
2004-02-01

Electronic imaging — Information stored electronically — Recommendations for trustworthiness and reliability

Images électroniques — Stockage électronique d'informations — Recommandations pour les informations de valeur et leur fiabilité

Reference number
ISO/TR 15801:2004(E)

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

Published in Switzerland

Contents

Page

Foreword vi

Introduction vii

1 Scope 1

2 Normative references 1

3 Terms and definitions 2

4 Information management policy 2
4.1 General 2
4.2 Information Management Policy Document 2
4.2.1 Contents 2
4.2.2 Information covered 3
4.2.3 Storage media 3
4.2.4 Image file formats 3
4.2.5 Standards related to information management 4
4.2.6 Retention schedule 4
4.2.7 Information management responsibilities 4
4.2.8 Compliance with policy 4

5 Duty of care 4
5.1 General 4
5.1.1 Controls 4
5.1.2 Separation of roles 5
5.2 Information security management 5
5.2.1 Information Security Policy 5
5.2.2 Risk assessment 6
5.2.3 Information security infrastructure 6
5.3 Business continuity planning 7
5.4 Consultations 7

6 Procedures and processes 8
6.1 General 8
6.2 Procedures Manual 8
6.2.1 Documentation 8
6.2.2 Content 8
6.2.3 Compliance with procedures 9
6.2.4 Updating and reviews 9
6.3 Document image capture 9
6.4 Document scanning procedures 10
6.4.1 General 10
6.4.2 Preparation of paper documents 10
6.4.3 Document batching 11
6.4.4 Photocopying 11
6.4.5 Scanning processes 12
6.4.6 Quality control 13
6.4.7 Rescanning 15
6.4.8 Image processing 15
6.5 Data capture 15
6.5.1 New data 15
6.5.2 Migration 16
6.6 Indexing 16
6.6.1 General 16
6.6.2 Manual indexing 16

6.6.3 Automatic indexing 16
6.6.4 Index storage 16
6.6.5 Index amendments 17
6.6.6 Index accuracy 17
6.7 Authenticated output procedures 17
6.8 File transmission 18
6.8.1 Intra-system data file transfer 18
6.8.2 External transmission of files 18
6.9 Document retention 19
6.10 Information destruction 20
6.11 Backup and system recovery 20
6.12 System maintenance 21
6.12.1 General 21
6.12.2 Scanning systems 21
6.13 Security and protection 21
6.13.1 Security procedures 21
6.13.2 Encryption keys and digital signatures 22
6.14 Use of contracted services 22
6.14.1 General 22
6.14.2 Procedural considerations 23
6.14.3 Transportation of documents 24
6.14.4 Use of trusted remote archives 24
6.15 Workflow 24
6.16 Date and time stamps 25
6.17 Version control 25
6.17.1 Information 25
6.17.2 Documentation 25
6.17.3 Procedures and processes 26
6.18 Maintenance of documentation 26

7 Enabling technologies 26
7.1 General 26
7.2 System Description Manual 27
7.3 Storage media and sub-system considerations 27
7.4 Access levels 28
7.5 System integrity checks 28
7.5.1 General 28
7.5.2 Digital and electronic signatures (including biometric signatures) 29
7.6 Image processing 29
7.7 Compression techniques 30
7.8 Form overlays and form removal 31
7.9 Environmental considerations 31
7.10 Migration 32
7.11 Information deletion and/or expungement 32

8 Audit trails 32
8.1 General 32
8.1.1 Audit trail data 32
8.1.2 Creation 33
8.1.3 Date and time 33
8.1.4 Storage 34
8.1.5 Access 34
8.1.6 Security and protection 34
8.2 System 35
8.2.1 General 35
8.2.2 Audit trail information 35
8.2.3 Migration and conversion 35
8.3 Stored information 35
8.3.1 General 35
8.3.2 Information capture 36
8.3.3 Batch information 37

8.3.4 Indexing 37
8.3.5 Change control 38
8.3.6 Digital signatures 38
8.3.7 Destruction of information 38
8.3.8 Workflow 38

Bibliography 39

Foreword

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

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

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

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

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

ISO/TR 15801 was prepared by Technical Committee ISO/TC 171, *Document management applications*, Subcommittee SC 3, *General issues*.

Introduction

Increasingly, information that has been created, captured and stored electronically is used as evidence of business activities. Such evidence might be required in contract discussions, or in courts of law. This Technical Report defines recommended practices for electronic storage of business or other information in image form. As such, complying with its recommendations is of value to organizations even when the trustworthiness of the stored information is not being challenged.

Users of this Technical Report should be aware that the implementation of these recommendations does not automatically ensure acceptability of the evidence encapsulated by the information. Where stored electronic information may be required in court, implementers of this Technical Report are advised to seek legal advice to ascertain the precise situation within their relevant legal environment.

This Technical Report describes means by which it may be demonstrated, at any time, that the contents of a specific electronic image file created or existing within a computer system have not changed since it was created within the system or imported into it. Where such a data file contains a digitized image of a physical source document, it will be possible to demonstrate that the digitized image is a true facsimile of that source document. The issue being addressed is essentially one of authentication.

Other versions of the information may legitimately develop; e.g. revision of a contract. In these cases the new versions are treated as new image files.

The same principle can be applied when a significant change is made to a document in a workflow environment.

This Technical Report describes procedures whereby an electronic copy may be demonstrated to be a true copy of the original, whether that original was itself an electronic data file or a physical source document.

The recommendations in this Technical Report are a mixture of items that are broad and general and items that are specific and detailed. Readers are advised to use this Technical Report in conjunction with other local sources, particularly with relevance to governmental and legal requirements in their respective jurisdictions.

Organizations that implement most of the recommendations described in this Technical Report will be in a good position to be able to demonstrate authenticity. However, there may be good economic reasons where a particular recommendation is not implemented. In such situations, the risk taken by such non-implementation decisions should be assessed.

Electronic imaging — Information stored electronically — Recommendations for trustworthiness and reliability

1 Scope

This Technical Report describes the implementation and operation of information management systems which store information electronically and where the issues of trustworthiness, reliability, authenticity and integrity are important. The whole life cycle of a stored electronic document is covered, from initial capture to eventual destruction.

This Technical Report is for use with any information management system, including traditional document imaging, workflow and COLD/ERM technologies, and using any type of electronic storage medium including WORM and rewritable technologies.

Image files may potentially contain any type of data: for example, correspondence, forms, drawings. This Technical Report covers all such image files, whether created and/or imported directly or through a network, from the time at which the system assumes control of the image file.

This Technical Report does not cover processes used to evaluate the authenticity of information prior to it being stored or imported into the system. However, it can be used to demonstrate that output from the system is a true reproduction of the original document.

Where in this document the term *system* is used, it should be taken as meaning the *information management system* that is being reviewed, unless otherwise stated.

2 Normative references

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

ISO 9000 (all parts), *Quality management and quality assurance standards*

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

ISO 12651:1999, *Electronic imaging — Vocabulary*

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

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 12651 and the following apply.

3.1
information type
groups of related documents

NOTE In specific applications, “groups” may be identified as “sets”, “files”, “collections” or other similar terms.

EXAMPLES Invoices, financial document, data sheets, correspondence.

4 Information management policy

4.1 General

Information is one of the most important assets that any organization has at its disposal. Everything an organization does involves using information in some way. The quantity of information can be vast, and there are many different ways of representing and storing it. The value of information used and the manner in which it is applied and moved within and between organizations may determine the success or failure of those organizations.

Information, like any other asset, needs to be classified, structured, validated, valued, secured, monitored, measured and managed efficiently and effectively.

This clause describes documentation that states the organization’s information management policy. Availability of this documentation will demonstrate that information management is part of normal business procedures.

Where a system stores records, compliance with ISO 15489 should be considered. Such compliance will ensure that many of the elements of this Technical Report will be implemented. Where a system stores information that may be used as evidence in court, your legal advisors should be consulted (see 5.4) to ensure that you comply with relevant legal or regulatory requirements. As legal and regulatory requirements vary from country to country (and sometimes within a country), the legal advice you obtain should cover all relevant jurisdictions.

4.2 Information Management Policy Document

4.2.1 Contents

An Information Management Policy Document (the Policy Document) should be produced, documenting the organization’s policy on information management and storage, as applicable to the information management system.

The Policy Document should contain sections which:

— specify what information is covered (see 4.2.2);

— state policy regarding storage media (see 4.2.3);

— state policy regarding image file formats and version control (see 4.2.4);

— state policy regarding relevant information management standards (see 4.2.5);

— define retention and destruction policies (see 4.2.6);

— define responsibilities for information management functions (see 4.2.7);

— define responsibilities for monitoring compliance with this policy (see 4.2.8).

The Policy Document should be approved by senior management of the organization, and should be reviewed at regular intervals.

Essential to this Technical Report is the agreement and implementation of a Retention Schedule for stored information. Where reference is made to the Policy Document in the rest of this Technical Report, the Retention Schedule is included in such a reference.

4.2.2 Information covered

In order to define the organization's information management policy, information should be grouped into *types*, the policy for all information within a type being consistent. For example information types may be specified either by reference to application (e.g. financial projections, invoices, customer address list), or to generic group (e.g. accounting data, customer documents, manufacturing documents).

During the drafting of the Policy Document, specific information may need to be moved from one type to another to ensure consistency of Policy within an information type.

The Policy Document should list all types of information which are to be stored. The Policy Document should include as an information type all documents produced in compliance with the Policy.

4.2.3 Storage media

Different types of media have different long-term storage characteristics. Most organizations will store information on a variety of media types: paper; microform; electronic (write-once and rewritable/erasable). In some applications, specific pieces of information may, throughout its retention period, be stored on different media types at different times.

The organization should have policies regarding the use of specific types of media for different information storage requirements (e.g. access requirements, retention periods, and security requirements). These policies should be detailed in the Policy Document.

NOTE In some countries, only certain media types can be used where stored information may be required as evidence. For example, in France, rewritable media cannot be used in evidential matters.

The media type on which each information type (see 4.2.2) may be stored should be specified.

It should be possible, where copies of image files exist, to be able to trace back to the earliest such files, in order to be able to determine that no changes have occurred to any purported copy. Note that, in the case of files that exist in different versions, each version should be treated as a new source or original file for the purposes of this Technical Report.

The policy for tracking copies of image files should be detailed in the Policy Document.

4.2.4 Image file formats

The Policy Document should contain details of the approved image file formats that may be used for each information type.

All information stored on a computer system requires software for retrieval and display. This software is subject to change, either by the implementation of new releases, or by changes to operating systems and/or hardware. By implementing a policy of approved storage formats, the necessary data migration or alternative procedures can be implemented satisfactorily to ensure long-term retrieval of the stored information.

Where compression techniques are available, policy on their use should be documented.

Where multiple versions of a document may be stored, a policy is required which ensures that all relevant versions are stored, and their relationship maintained. The Policy Document should contain details of policy on the storage of versions of documents.

For additional information on this, see 6.5.2, 7.10, and 8.2.3.

4.2.5 Standards related to information management

Where the organization operates a quality management system (such as the ISO 9000 series), whose scope includes part or all of the information management system, then all relevant procedural documentation should be included in the quality system.

Where National or International regulatory requirements are mandatory, or where National or International Standards are applicable, they should be complied with.

4.2.6 Retention schedule

A Retention Schedule should be established for each information type.

Retention periods should be agreed by all relevant departments and personnel within the organization.

Retention periods should be agreed upon after taking relevant advice to ensure that legal or regulatory issues, or both, are resolved.

All relevant system and procedural documentation that is produced should be covered by the Retention Schedule.

The Retention Schedule should include the organization's policy for its periodic review.

The Retention Schedule should include the organization's policy for the controlled destruction of information.

4.2.7 Information management responsibilities

Individual or job function responsibilities for the Policy Document should be defined in the Policy Document.

Individual or job function responsibilities for each information type should be identified and included in the Policy Document.

Individual or job function responsibilities should include the need to seek relevant advice when creating or updating the contents of the Policy Document.

4.2.8 Compliance with policy

Where it is important that compliance with the Policy Document can be demonstrated, the individual or job function responsibilities for obtaining and maintaining such compliance should be identified and defined.

5 Duty of care

5.1 General

5.1.1 Controls

It is essential that the organization is aware of the value of information that it stores, and executes its responsibility under the duty of care principle.

To fulfil this objective, the organization needs to:

— establish a chain of accountability and assign responsibility for activities involving management of electronic information at all levels;

— be aware of legislative and regulatory bodies pertinent to its business;

— keep abreast of technical, procedural, regulatory and legislative developments by maintaining contact with the appropriate bodies and organizations;

— implement an Information Security Policy.

5.1.2 Separation of roles

The separation of roles is a fundamental aspect of duty of care. It provides a check on errors and on the deliberate falsification of records (in this respect separation of roles is particularly important in systems where there is risk of fraud or other malicious action).

There are several aspects of information management where a separation of roles is considered:

— input reconciliation (see 6.4.3);

— quality control (see 6.4.6);

— data entry (see 6.6.2);

— information deletion (see 6.9);

— information security (see 5.2).

It is also important to ensure that the physical and managerial separations that exist around a system are mirrored by the logical access controls within it.

The separation of roles between initial operations and checking should be reviewed and implemented where appropriate.

5.2 Information security management

5.2.1 Information Security Policy

All information, irrespective of the media on which it is stored, is vulnerable to loss or change, whether accidental or malicious. To protect information stored electronically, security measures need to be developed and implemented to reduce the risk of a successful challenge to its authenticity. These security measures need to be aligned to any information classification categories that are used.

Traditionally, information security is considered a matter of confidentiality, to ensure that information is not accessible outside the requirements of the organization. However, whilst this is important (in some cases vital) to the operation of the organization, it is not the most important security issue relevant to this Technical Report.

A key objective of the Information Security Policy is to ensure the protection of the integrity of stored information. When developing security measures, it is necessary to compare the risk of integrity being compromised with the cost of implementation of such measures. Security measures need to include backup and other copies of stored information, as their integrity is of importance in circumstances where they have been used as replacements for live data.

Also of importance is availability. In some cases, it may be necessary to be able to demonstrate that all information on a specific topic is available for review at any time. In this category, topics such as indexing accuracy and business continuity planning are key.

Security is not singularly a concern of computer systems. Security and availability of the operating environment (including buildings, temperature controls, network links, etc) and the auditable implementation of procedures by all staff are both key elements.

The organization should adopt an Information Security Policy, covering all elements of the information management system.

Where the organization has an Information Security Policy for other systems, then the use of the information management system should be incorporated within its scope.

The Information Security Policy Document should contain, as a minimum:

— scope of policy;

— statement of management objectives in respect of security;

— specific policy statements;

— requirements for different information classification categories;

— definition and allocation of information security responsibilities;

— policy for dealing with breaches of security;

— policy regarding compliance with relevant standards.

The Information Security Policy Document should be approved by the organization's senior management. That approval should be documented.

The organization should agree and document appropriate levels of security for managing its information, in compliance with its Information Security Policy Document.

5.2.2 Risk assessment

Security measures are often developed using an *ad hoc* approach, reacting to security incidents or to available computer software tools. Such procedures frequently leave gaps in security, which are only filled at some later date. A more structured approach is to review the information assets of the organization, and assign risk factors (based on asset value, system vulnerability and likelihood of attack). An Information Security Policy can then be produced and approved, against which security measures can be audited.

The organization should undertake an information security risk analysis, and document the results obtained.

Of particular importance are the security measures implemented to control the information storage media, both the live media and the backup media. The risk analysis needs to include vulnerability risk factors consistent with the type of media being used (e.g. WORM or rewritable).

Where different types of storage media are used, their impact on the risk analysis results should be reviewed.

Once the risk analysis has been completed, it needs to be acted upon as part of a review of implemented security measures. Factors such as the balance between the cost of implementation, security achieved and risk evaluation need to be taken into consideration during the review process.

Based on the results of the risk analysis, existing security measures should be reviewed for effectiveness.

Where the review indicates that changes to security procedures are appropriate, the changes should be implemented.

5.2.3 Information security infrastructure

A management framework should be established to initiate and control the implementation of information security within the organization. The framework should have as its objectives:

— approval and review of the Information Security Policy;

— monitoring of threats to information security;

— monitoring and review of security breaches;

— approval of major initiatives to enhance information security.

5.3 Business continuity planning

From time to time, problems arise with information management systems that require emergency procedures to be implemented, to recover from the problem. Such procedures may involve the temporary use of additional or third-party resources. In order to ensure that the integrity of information is not compromised during these operations, an agreed and approved Business Continuity Plan (sometimes known as a Disaster Recovery Plan) may be implemented.

Procedures to be used in cases of major equipment, environmental or personnel failure should be developed, tested and maintained. Such procedures should ensure that the integrity of stored information is not compromised during their implementation.

5.4 Consultations

The implications of using electronic information management systems may be significant to other organizations, such as:

— regulatory bodies;

— government bodies;

— external audit bodies;

— legal advisors (such as the organization's lawyers).

The organization should determine the levels to which consultations should be made.

These levels may include the following:

— national and international law;

— industry sector;

— community;

— organization;

— department;

— individual.

The organization should consult with relevant organizations prior to implementing the Information Management Policy Document.

These consultations may include the following topics:

— legal issues;

— government regulations;

— financial regulations (such as payment of taxes);

— special regulations (applicable to particular sectors).

The results of all consultations, including actions agreed, planned or implemented, should be referenced or included in the Policy Document.

Where appropriate regulations and/or laws exist, they should be complied with.

The Policy Document should state whether all or part of any relevant national or international standards should be complied with.

Where the organization complies with relevant national or international standards, such compliance should include the information management system.

6 Procedures and processes

6.1 General

This clause deals with the operating procedures that the organization needs to review and, where appropriate, implement. It refers to a single information management system. Where the organization has multiple information management systems, procedures related to each such system can be produced separately or in combination.

6.2 Procedures Manual

6.2.1 Documentation

The organization should maintain a Procedures Manual for each information management system, describing all procedures related to the operation and use of the system, including input to, operation of and output from the system.

Where in this section documentation is required, this documentation can either be included in the Procedures Manual, or referenced by it. This Manual may include references to other controlled documentation as appropriate.

The relevant procedures detailed in, or referenced by, the Procedures Manual should be readily accessible to all appropriate users of the system.

6.2.2 Content

The Procedures Manual should include the topics listed below:

— document capture (see 6.3);

— document scanning (see 6.4);

— data capture (see 6.5);

— indexing (see 6.6);

— authenticated output procedures (see 6.7);

— file transmission (see 6.8);

— information retention (see 6.9);

— information destruction (see 6.10);

— backup and system recovery (see 6.11);

— system maintenance (see 6.12);

— security and protection (see 6.13);

— use of contracted services (see 6.14);

— workflow (see 6.15);

— date and time stamps (see 6.16);

— version control (see 6.17);

— maintenance of documentation (see 6.18).

For convenience, the Procedures Manual may be maintained as a number of separate physical documents, relating to different information management areas.

Where the organization has multiple information management systems, the documentation may comprise a single Procedures Manual or multiple Procedures Manuals.

6.2.3 Compliance with procedures

In order to be able to comply with the procedures detailed in the Procedures Manual, staff needs to be aware of them, and have the ability to follow them. This situation is frequently achieved by training, either by specific courses or during day-to-day working.

Procedures should be implemented which ensure that all staff who operate the system adhere to requirements.

6.2.4 Updating and reviews

It is important to ensure that the procedures implemented at any time during the storage life of any specific piece of information can be determined. This is achieved by ensuring that the Procedures Manual is kept up to date, and that all previous versions are kept in compliance with the Policy Document.

Any changes to operational procedures should be documented. This documentation should include details of any change control procedures used, and procedures to ensure that the new procedures are implemented.

Where changes are being implemented, they should be checked to ensure that operational requirements and the requirements of the Policy Document are not compromised.

Superseded versions of the Procedures Manual should be kept in compliance with the Policy Document.

To confirm that documentation is up to date, regular reviews are necessary. Such reviews may also be necessary where legal or regulatory changes are relevant.

A review should be carried out at least annually to ensure that any changes to procedures or technology are reflected in the Procedures Manual.

The results of periodic reviews should be documented and approved by the person responsible for the operation of the appropriate part of the system.

6.3 Document image capture

Where the information management system is used for storing document images, the procedures involved in the capture of those images should be documented.

These procedures may include:

— document preparation;

— document batching;

— photocopying;

— scanning;

— image quality control.

Source documents may include paper documents or microform documents.

Subclause 6.4 contains further details on the procedures relevant to document scanning.

6.4 Document scanning procedures

6.4.1 General

This subclause includes recommendations relating to the procedures relevant to document image capture. Recommendations in this subclause are for users whose information management systems include the capture and storage of document images by the use of scanners. These recommendations cover procedures for:

— preparation of documents;

— document batching;

— photocopying;

— scanning;

— image processing.

6.4.2 Preparation of paper documents

All paper documents need to be examined prior to the scanning process, to ensure that a successful image is obtained. Attributes such as paper size, weight and binding, paper and print colour can all affect the physical scanning process.

Documents should be examined prior to the scanning process, to ensure their suitability for scanning. Procedures for this examination process should be documented.

Factors such as their physical state (thin paper, creased, stapled, etc.), and the attributes of the information (black-and-white, colour, tonal range, etc.) should be considered.

Where documents are found which are unlikely to be accepted by the scanner, there are a number of techniques that can be used. For example, the original could be photocopied, or transparent wallets could be used.

Procedures to be followed for documents that may cause scanning difficulties should be documented.

When removing staples, clips or other document bindings, ensure that no damage is caused to the original that may affect the capture of the information from the document.

Where a source document has physical attachments, for example, stick-on notes, the system should provide facilities for distinguishing these from the document to which they are attached.

This might be achieved, for example, by capturing a separate image of the attachment, with appropriate data to associate it with the source page. If only a single image is captured with the attachment in place, the data might record the fact that there is an attachment. Where there is a risk that an attachment might obscure, or be considered to obscure, information on the source document, it might be preferable to ensure that an image of the source document without the attachment is captured.

Where a source document has physical amendments, for example, white opaque paint, the system should ensure that the presence of such amendments is noted.

Procedures to be followed when scanning multi-page documents bound together with staples or clips should be documented.

All pages of multi-page documents should be kept together and in the appropriate order before, during and after scanning.

6.4.3 Document batching

Wherever possible, documents should be grouped into batches.

This makes it easier to control documents, and to be able to perform quality control and other procedures on a sampling basis.

The definition of batch size should be decided on the basis of convenience.

The number of documents in a batch will be application dependent. For example, if the documents are in file covers, and the average number of documents per file cover is relatively large, for example 100 pages, then the documents in a single file cover may constitute a batch. If the file covers contain relatively few documents, for example on average 10 pages, then a batch may consist of documents from more than one file cover. If the documents are on roll microfilm, the film roll may be a batch.

Choose the batch size so that it is not bigger than can be easily managed, nor so small that checking quality by sampling on a batch basis would result in significant process inefficiencies. Sample size may need to be determined using statistical sampling techniques.

For some applications, a batch may not be easily defined. In these cases, a batch may be defined as *those documents input during a specified time period*. Thus, for example, a batch could be all documents input during an hour or a day.

For some applications (especially where workflow is implemented) where batching cannot be applied, alternative methods for ensuring that all documents are scanned should be established. Such techniques may include the marking of documents after scanning, or additional checking of images against the paper originals.

6.4.4 Photocopying

Where source documents are photocopied prior to the scanning process, the procedures used should be documented in the Procedures Manual.

It may be helpful for some documents to be photocopied prior to being scanned. Such documents include:

— documents that may be adversely affected by the scanning process, such as damaged or delicate documents;

— documents where there are substantial contrast or density variations over the area of the original, and where photocopying demonstrably improves the image quality;

— documents containing paper or ink colours that do not produce legible scanned images;

NOTE 1 Photocopiers and scanners may respond differently to different colours, and it is only in exceptional cases that the technique of photocopying prior to scanning does not produce satisfactory results.

— folded documents that are too large to be scanned as a single full-sized image.

NOTE 2 Photo-reductions may be made which are then scanned, and/or multiple scanned images may be captured from the original or from photocopies thereof.

Photocopies should be examined to ensure that there is no significant loss of information during this process.

Additional quality control procedures should be adopted to ensure that no significant information is lost in the scanning of photocopied original documents.

If photo-reductions are made, checks should be made to ensure that there is no significant loss of detail in the scanned images compared to the original caused by the effective resolution of the image (compared to the original) being reduced.

If multiple images are captured, these should be overlapped to ensure that there is no significant loss of information at the edges between adjoining images.

Where an image was made from a photocopy, it should be clear to a user of the image that this was the case. It should also be clear whether the photocopy was made from the source document during document preparation or whether the source document was known to be a photocopy.

This is to ensure that an image may be correctly identified as a true facsimile of a source document, even if an intermediate photocopy has been taken as part of the preparation procedures, and to distinguish such images from images of photocopies made under unknown conditions.

This may be done, for example during document preparation, by stamping or marking the document as a *photocopy* or *original photocopy*, or by electronically marking the image as having been captured from a photocopy, distinguishing between photocopies made during document preparation and source documents which are known photocopies.

Procedures to be used where it is not known whether a document is an original or a photocopy should be documented.

6.4.5 Scanning processes

Details of procedures used in document scanning should be included in the Procedures Manual.

Any variations in scanning procedures due to the type of document being scanned should be detailed in the Procedures Manual.

Such changes may apply to, for example, double-sided versus single-sided documents; colour versus black-and-white images.

Procedures should ensure that all documents in a batch are fully scanned; no document should be left unscanned.

To check that all documents have been scanned, the count of captured documents may be compared with the number of documents in a batch. Where batching is not used, alternative procedures for ensuring that all documents are scanned may be needed.

Where it is important that all pages in a multi-page document are scanned, procedures which ensure this should be implemented.

The count of captured images per document may be compared with the number of pages (i.e. sides) in each document, taking into consideration any blank page (or other) removal processes. However, errors in manually counting physical documents and the pages therein may make such a process ineffectual. It may be satisfactory to implement procedures whereby the probability and risk of any document not being scanned is acceptably small. This risk should be evaluated and, where necessary, procedures should be reviewed against this risk.

Many scanners have automatic document feeders that can reliably detect document misfeeds, therefore minimizing the risk that a document may pass through the scanner without being scanned. If such devices are not used, the procedures are required to ensure that the scanner operator has to manually handle every document in order to reduce the probability that any document might not be scanned.

Where it is crucial to ensure that every sheet is scanned, users should consider counting or pre-indexing the documents in order to capture accurately the number of pages per document or batch of documents.

Using a double-entry technique can provide extremely high accuracy in the number of pages. This data may subsequently be compared with the scanned page count; any shortfall will indicate either that more than one page has been fed at once or that a page has been misplaced between pre-indexing and scanning.

If a simplex scanner (i.e. one that scans only one side of a document at a time) is used to scan double-sided documents, care should be taken to ensure that every double-sided document is reversed and the other side scanned.

If a large source document is scanned in sections, so that multiple images are captured, these sections should be overlapped to ensure there is no loss of information at the edges between adjoining images.

The scanning system should enable each document to be uniquely identified, in such a way that its identity cannot be changed or removed, except as permitted in the subclause of this Technical Report on expungement (see 7.11).

This unique identity could be a system-generated sequence number, which may be used for internal control purposes only.

6.4.6 Quality control

6.4.6.1 Sample set

Procedures are required which reduce the risk of scanned images being of unsatisfactory quality. It will be easier to demonstrate authenticity if it can be shown that the images are of good quality, and that the scanner was working to agreed standards at the time of scanning.

A sample set of source documents, or of documents equivalent in characteristics to the source documents, should be assembled for the purposes of evaluating scanner results against agreed quality control criteria. Documents in the sample set should be representative of the complete set of documents that are to be scanned. Documents in the sample set should include examples of source documents whose quality is poor relative to those of the majority of the documents.

Quality control criteria may cover:

— overall legibility;

— smallest detail legibly captured (e.g. smallest type size for text; clarity of punctuation marks, including decimal points);

— completeness of detail (e.g. acceptability of broken characters, missing segments of lines);

— dimensional accuracy compared with the original;

— scanner-generated speckle (i.e. speckle not present on the original);

— completeness of overall image area (i.e. missing information at the edges of the image area);

— density of solid black areas;

— colour fidelity.

Quality control criteria for image quality should be realistic given the nature of the source material and the characteristics of the scanning equipment.

Quality control criteria should be documented for scanned image quality. The criteria should be agreed by all parties whose use of images is likely to be affected by image quality, including internal and external users.

Quality control criteria should be based upon the sample set of documents.

6.4.6.2 Evaluating image quality

Procedures that specify the process used for evaluating image quality on a day-to-day basis should be documented.

Image quality evaluation procedures should include details of the evaluation of results, including the characteristics of the image retrieval device.

Take care when evaluating the results of a quality control procedure. Results obtained may depend upon the specific output device (e.g. monitor or printer).

If a printer is to be used for quality control procedures, the printer resolution should be equal to or greater than the resolution of the scanned images.

The printer should be capable of accurate reproduction of grey scale or colour in applications where this is relevant.

Where grey scale or colour reproduction is relevant, the accuracy of rendition of grey scale or colour should be evaluated.

Where dimensional accuracy is important, procedures should be documented for checking that dimensional information is reproduced within tolerance. This may involve, for example, checking that the nominal resolution of the scanner is accurate, so that the dimensions in the digital image may be determined by counting the number of pixels between specific points in the image.

If the scanner operator checks the quality of images during the scanning procedures, a second quality control procedure should be undertaken by personnel other than those responsible for the scanning. This second quality check may involve statistical sampling techniques.

Quality control procedures should be related to the batch process (if used) as defined earlier, enabling acceptance or rejection of such a batch independently of any other batch.

The results of all quality control checks should be stored in the Quality Control Log (which may be created manually or automatically).

In workflow environments where every document is viewed within a workflow process, and activities explicitly check images for quality and reject unacceptable ones, then these activities may be deemed to be a quality control process.

Where the quality control procedures involve sampling of the scanned images and any related data (such as notes), the proportion sampled need not be fixed but may vary from time to time depending on the frequency of problems encountered or the nature of the source material. Where appropriate, statistical sampling techniques should be used to determine percentage of scanned images to be checked.

It will not normally be practicable to check all processed material and generally only a proportion of the material processed will be checked. For example, when starting scanning initially a relatively large sample may be selected (e.g. 20 %), which may be reduced (e.g. to 10 % or even 5 %) as the consistency of meeting the required quality standards can be demonstrated.

Where quality control consists of sampling scanned images, the frequency of sampling should be documented.

6.4.6.3 Checking scanner performance

Scanner performance checks should be used periodically to monitor the system, to check that it is within agreed tolerances.

Hard copy prints may be made of the scanned images of the test targets and compared with the test targets themselves to determine whether the quality criteria are met, as described in the procedures.

Test targets allow objective assessment and measurement of scanner performance. Regular use can show whether the scanner is performing consistently and in accordance with its specification. The test target given in ISO 12653-2 may be used for this assessment.

The frequency of scanner performance checks should be dependent upon system usage, and related to expected deterioration in system performance. This may require recommendations from the system supplier and also experience in the use of the system. Initially, it may be appropriate to scan a test target every few thousand pages scanned.

If double-sided (duplex) scanners are used, double-sided test targets should preferably be used. Single-sided test targets should only be used with duplex scanners if double-sided test targets cannot be obtained.

Test targets are not representative of the documents actually being scanned and are not to be regarded as a substitute for the sample set of documents.

6.4.7 Rescanning

Procedures for rescanning documents should be documented. Such rescanning may be required if an original image has been rejected, owing to poor quality or other factors.

Procedures should be implemented to ensure that images resulting from rescanning replace the original image, and that batch numbering and audit trail procedures are not compromised.

6.4.8 Image processing

Image processing techniques used to improve the quality of an image should be described in the Procedures Manual.

Where procedures involve frequent change from document to document, specific settings need not be stored, but should be referenced in the System Description Manual

Where operator-controlled facilities are available for use, details of which facilities are used for a particular document should be documented.

6.5 Data capture

6.5.1 New data

Data (for example for the creation of index or other reference information) may be captured from existing documents and entered into a computer in a number of ways, including manual (i.e. direct keyboard entry), automated (e.g. bar code reading, Optical Mark Reading (OMR), OCR/ICR), or semi-automated (e.g. where data captured automatically, e.g. by OCR, is confirmed by manual re-entry). In each case, the issue is to convey confidence that the correct data has been captured. In practice it can be difficult, if not impossible, to ensure 100 % accuracy in captured data, and the user has to assess the risk associated with the existence of errors.

Where external data is captured for entry into the system, required quality levels should be specified. These quality levels should cover accuracy and completeness of captured data.

The specified accuracy levels may vary depending on the application and the importance of each particular data item.

Procedures should be defined for checking that the accuracy levels are maintained. These procedures will typically be based on random or quasi-random sampling of batches of captured data, with comparison against the source material. Batches that fail to meet the required accuracy levels will generally be reprocessed and the results checked again to ensure that the required accuracy levels are maintained.

Records should be kept of the results of all accuracy checking.

6.5.2 Migration

Where data is being received from another system (or part of a system), as part of a system migration process, then procedures and processes need to be established, implemented and documented for this process. Where information may be lost (for example by change of format or resolution) as part of a transfer between systems, such loss should be documented.

Where data files and any associated metadata are being received from a compliant system as part of a migration process, procedures and processes should be documented.

6.6 Indexing

6.6.1 General

Indexing is a vital part of the process of storing information on electronic media, as it allows access to the relevant information. Where indexing information is lost, then the stored information may also be lost.

Indexing can be either automatic (i.e. performed by the system without operator intervention), or manual. If manual indexing is performed, it is important to ensure that the documented procedures are followed.

Some systems allow partial index information to be stored when the information is captured. This may then be combined with additional manual index entries at a later time.

Procedures and rules for indexing stored information should be documented.

6.6.2 Manual indexing

Manual indexing involves the visual examination of information being captured by the system, either prior to its capture or as part of post-capture processes.

Staff involved in manual indexing should receive specialist training, to maximize accuracy. Indexing training requirements and procedures should be documented.

6.6.3 Automatic indexing

Automatic indexing may be effected by, for example, the reading of bar codes, or the use of OCR/ICR techniques. Where automatic indexing is used, procedures to check and amend inaccurate index data should be documented.

6.6.4 Index storage

Index data should be retained for at least as long as the information to which they relate is retained.

Some systems require database indexes to be rebuilt periodically, typically to improve database performance. Procedures for rebuilding indexes should be documented.

6.6.5 Index amendments

Indexing processes may include procedures for the detection of missing information. Indexing from displayed information will not detect missing material unless the displayed information is checked against the originals, or there is a defined sequence of information (for example by sequential numbering).

Procedures for the amendment and/or correction of indexing data should be documented. If an index entry is amended, details of index content before and after the change may need to be retained.

Where an index entry relates to deleted or expunged information, this status should be stored.

Where deletion or expungement of stored information, by the amendment or deletion of index entries, may be required to comply with legal or regulatory requirements, procedures to be followed should be documented.

6.6.6 Index accuracy

Index data for scanned images may be inaccurate. While accurate indexing will facilitate the retrieval of stored information, the authenticity of that information may be demonstrated if its relevance and completeness can be indicated from the accuracy of the relevant index data. Conversely, inaccurate index data may result in the user being unable to retrieve relevant information, or retrieving irrelevant information.

Index data accuracy criteria may vary depending upon the application. In some cases, the accuracy may be defined as the maximum acceptable number of characters in error per 1 000 characters captured (or percentage equivalent). In other cases, the accuracy may be defined as the maximum acceptable number of words (or similar cluster of characters, for example a customer or part number) containing any error (whether of one or more characters).

Criteria for index data accuracy levels ought to be realistic, given the method used for index data capture, the typical random error rates achieved by data entry personnel, and the legibility of the source material. These accuracy levels may vary depending upon the type of information being indexed.

Where manual or automatic indexing is undertaken, accuracy levels should be agreed and documented. Procedures for index data accuracy checking should be documented.

6.7 Authenticated output procedures

Output from electronic storage systems, either in the form of paper copies or as electronic files on appropriate storage media, may need to be produced for use as documentary evidence. Generally, these copies need to be authenticated as true copies of the original, to reduce the likelihood of rejection or challenge.

Procedures for the creation of authenticated copies should be documented.

Such procedures may, for example, require the use of standard system features for copying, and written confirmation by an authorized person that the copying process has been conducted correctly. The procedures may specify how authenticated copies are subsequently to be handled. The procedures may refer to audit trail data as a confirmation of the processes that occurred during copying.

Where a physical document is produced as part of the output, the procedures should include the use of an authorized signature or other procedure to authenticate this copy document.

It is important that the nature and extent of any changes introduced by the retrieval facilities are understood and their relevance assessed. What is acceptable in normal usage may be unacceptable in other circumstances. For example:

— rendering a coloured image in monochrome may be acceptable in situations where the colour is irrelevant; but in other situations the colour may be vital, necessitating a different retrieval facility;

— viewing an image at a lower resolution than that used in scanning the original document may be acceptable in routine retrievals, but the fine detail which is thereby lost may be important in other situations where, for example, it might have forensic significance;

— where there is not an exact match between the resolution of a scanned image and the retrieval device, the dimensional accuracy of the reproduction may be lost;

— where a stored data file is normally converted to another format for display or printing, information may be lost or presented in a different form, caused by loss of detail or layout differences. These differences may be unacceptable for disclosure, and in these cases different retrieval facilities may be required which do not involve conversion.

If the system facilities used to retrieve, display and/or print stored information do not maintain the layout (e.g. font, pagination) of the original, information retrieval characteristics should be agreed and documented.

6.8 File transmission

6.8.1 Intra-system data file transfer

6.8.1.1 General

Intra-system file transfers are those that take place within the system as defined in 7.2 of this Technical Report. Intra-system file transfers include:

— local area network transmissions;

— movement between storage sub-systems under system control, e.g. in a hierarchical storage management system, or between cache and magnetic disk;

— transfer between storage sub-systems under operator control.

In such transfers, the procedures, both electronic and manual, are under the control of the organization.

Procedures and processes should be implemented to ensure that the integrity of files transferred within the system is not compromised.

File transfers from one device to another should be controlled by the application software.

Where additional security measures are required, the use of digital signatures should be considered.

NOTE This subclause is not applicable to the requirement for file migration, where the media type and/or format of the data file may change for technology migration reasons. See 7.10.

6.8.1.2 Local area network transmission

In some applications, files may be transferred under operator control from one storage device to another using a local area network as defined in 7.2. Local area networks may include connections between remote locations using fixed lines. For dial-up or other connections, see 6.8.2.

Where files are transferred via a local area network, procedures and processes should be implemented to ensure that the integrity of files transferred is not compromised.

Where files are transferred between remote locations via fixed (e.g. leased) communications lines, procedures and processes should be implemented to ensure that the integrity of files transferred is not compromised.

6.8.2 External transmission of files

This subclause deals with files transmitted between one system and another via external, wide area, communications systems. Such systems are external to the system described in Clause 7. The sending and receiving systems are remote from each other and may be within the same or different organizations; in either case another party provides the transmission service.

The communications system may involve real-time transmission or deferred (store and forward) transmission such as occurs in e-mail services.

This Technical Report is concerned with the integrity of image files that has been transmitted to another party, and with the integrity of image files received from another party; this Technical Report is not directly concerned with the transmission service. By following the recommendations in this Technical Report, users can show that a copy of a file which was transmitted at some previous time to another party has not been altered since that time, and that a file received at some previous time via a transmission from another party has not been altered since the time of receipt.

File transfers from one device to another should be controlled by the application software.

Where a file is copied to another party via a transmission, the original file should be stored within the system.

The date and time of any file transmission should be stored as part of the audit trail.

Where a file is received from another party via a transmission, that file should be stored within the system.

The date and time of any file receipt should be stored as part of the audit trail.

Differences between sent and received files might be caused by errors in transmission or by deliberate alteration of one file or another. Demonstrating that a received and a sent file contain identical data is no different from demonstrating that any two copies are equivalent. The primary need is to show which file is the source, and which file is the copy; i.e. which file existed first. In some instances, this requirement can be met by comparing the times at which the two files were stored. If system time clocks are accurate (and bearing in mind differences in time zones), a received file should have been stored later than that at which the source file was transmitted. Thus, the issue becomes one of being able to demonstrate the reliability and accuracy of the timings of the two events.

Electronic/digital signatures, for example, may be used to permit confirmation that a received file or message is exactly the same as was sent, and to confirm the identity of the sender.

Additional procedures (outside the scope of this Technical Report) may be adopted for security or other reasons, e.g. to prevent unauthorized disclosure of the information contained within a file.

Where it is important to be able to demonstrate that a file has been delivered, the sender may require that the receiving system transmit back to the sender a confirmation of receipt, which should include the transmission identifier and the date and time of receipt.

If these procedures are followed, then the risk is reduced that a file has been modified, or has been sent from someone other than the identified sender.

The level of security risk being taken during an external file transfer should be assessed, to ensure compliance with the requirements of the Information Security Policy.

6.9 Document retention

Where source documents are scanned, and the Information Management Policy Document states that it is general policy to destroy a specific type of source document, there are some instances in which an exception applies and the source document ought to be retained.

Procedures that identify specific source documents that need to be retained should be documented.

Circumstances where this may be required include:

— where the source document is of poor quality, so that a legible image cannot be obtained;

— the source document may be kept to reduce the possibility of it being suggested that the image was deliberately made illegible; this also avoids any risk of rejection of an image on the grounds that it is not a facsimile of the source document;

— alternatively, a note may be stored which states that the original source document was of poor quality, and includes details of any visible information that needs to be stored;

— where a source document contains physical amendments or annotations that cannot be identified as such on the scanned image;

— a separate record that *physical amendments or annotations were present on the original document*, plus details of what the physical amendments were, may be sufficient;

— where fraud has been identified, or where litigation is envisaged or ongoing;

— documents of high value, such as the signed original of a large contract.

Procedures for the identification of information for which fraud has been identified, or for which litigation is envisaged or ongoing should be documented. Such procedures should include the suspension of document destruction policies for this information.

6.10 Information destruction

Procedures for the destruction or disposal of information at the end of the retention period should be documented.

These procedures should incorporate security precautions appropriate to the sensitivity of the information being destroyed.

No source documents should be destroyed until the images have been successfully written to storage and appropriate backup procedures have been completed.

6.11 Backup and system recovery

Effective procedures for the backup of files provide sufficient up-to-date copies to be used in the event of loss or corruption of part or all of the live data. It is vital that backup data includes all associated information (such as index files, audit trails), so that a complete new system can be built in the event of a total loss of the original system.

System recovery procedures also need to be documented, to demonstrate that such procedures are controlled and tested for reliability.

Issues surrounding the security of backup data may be important in the event of a dispute over authenticity. It may be argued that backup media had been compromised, and then used to recover from an information loss, thus affecting the authenticity of stored information. In some cases, the availability of backup data which has been in secure storage, to be used only in the event of a challenge to the authenticity of the live data, can be used to enable the demonstration of authenticity of the stored information.

Facilities on the system should allow for the backup and verification of all files and associated information, including audit trails, at regular intervals.

There should be information kept in the system audit trail of all backup activity, which should include details of any problems incurred during the procedure.

The procedures should include the secure off-site storage of these backups.

If the structure of the files held on a backup is different to that of the originals, the structure of the backup files should be detailed in the Systems Description Manual.

The audit trail should detail all file recovery activities, and include a description of any problems experienced during the recovery procedures.

Procedures for checking that file integrity has not been compromised after a restore should be documented.

Where backup data is used to recover from a system failure, procedures should be documented to ensure that file integrity has not been compromised.

Media used for backups do not necessarily provide permanent storage conditions. Media suppliers usually provide information regarding recommended testing frequency. Alternatively, if such specific information is not available, general recommendations can often be found in national or international standards.

Testing media on the same hardware each time is no guarantee that the media can be read on other devices, even of the same supplier and model type. Backups are of no value if the only hardware that can read them is lost.

Backup media should be tested at regular intervals, using a variety of hardware to read the media.

6.12 System maintenance

6.12.1 General

The information management system should be maintained and corrective maintenance carried out only by qualified personnel, to ensure that its performance does not deteriorate to such an extent that the integrity of the data captured or created by or stored within it is affected.

For example, it is of specific importance in a document scanning system that it is maintained in accordance with the manufacturer's specifications, in order that image quality is maintained.

Preventative maintenance should be carried out regularly, in accordance with the supplier's recommendations.

Procedures used for preventative maintenance should be documented.

These procedures may be performed by system operators, or by specialized service personnel.

A Maintenance Log should be kept, stating the preventative and corrective maintenance procedures completed.

Procedures to control the use of system maintenance hardware and/or software that can bypass system access controls should be documented. Access to such tools and facilities should be strictly controlled and monitored.

Information regarding system downtime, and details of action taken, should be stored in the Maintenance Log.

6.12.2 Scanning systems

Where document scanning is implemented, procedures described under the quality control section should be used to check that a scanning system continues to produce the output quality required of the system after the maintenance procedures have been completed.

These test results will serve to confirm, at any later date, that any poor quality images were not due to malfunction of the system. If there is any deterioration in the output quality, appropriate corrective maintenance is necessary.

6.13 Security and protection

6.13.1 Security procedures

Security guidelines that are applicable to the organization and application concerned should be implemented. Such guidelines, for example, might exist in company policies or practice, sector-specific guidance (e.g. financial, medical), National or International Standards, or as legal requirements.

In the absence of internal guidelines, published documents may provide comprehensive sets of information security guidelines that are designed to meet the organization's needs. They might provide an adequate basis for the creation of guidelines that would meet the organization's requirements. Some organizations may consider the adoption of externally accredited security schemes as additional confirmation of compliance with their Security Policy.

Procedures implemented in accordance with the organization's Information Security Policy should be documented.

To control access to the various levels of the system (e.g. manager, data input, and retrieval), a secure access control system should be implemented.

The accommodation and operating environment for information management systems and for the storage, labelling, handling, transportation and maintenance of storage media should be in accordance with suppliers' recommendations and/or relevant National or International Standards.

The central part of the system (including file servers, storage etc.) should be installed in secure areas (as defined in the organization's security procedures), with documented restricted access.

6.13.2 Encryption keys and digital signatures

Encryption techniques may be used to improve the security and integrity of stored data. A complete electronic file may be encrypted so that the information it contains cannot be retrieved without the use of an encryption key. Encryption is a complex topic, and one that is constantly changing. Readers should refer to authoritative publications on this topic for detailed information.

Digital signatures consist of data which, when appended to an electronic file, enable the user of the file to authenticate its origin and integrity. The digital signature data can be applied and checked by the application of public and private keys. The use of digital signatures does not imply that the file itself has to be encrypted. In many cases the file will be unencrypted; the digital signature serves to demonstrate whether the file contents have been tampered with and whether the file was signed by the purported signatory.

Where encryption or digital signatures are used, keys should be kept securely and should not be available except to those authorized as responsible for activities requiring access to the keys.

Procedures should be implemented for encryption key allocation and management and for certificate management where digital signatures are used.

Where encryption and digital signatures are used, and additional benefits can be obtained from third-party key management/recovery and key escrow services, their use should be considered.

The person who originally was responsible for managing the keys and certificates securely within the organization may no longer be employed, so procedures are required to ensure their continued availability.

In some countries, the use of encryption techniques is either restricted or illegal. The use of digital signatures may be permitted even though file encryption may be restricted or illegal. Particular care should be taken to meet legal requirements related to encryption in countries where information is stored, or through which it is transmitted.

6.14 Use of contracted services

6.14.1 General

Specialist service providers are often used for document scanning, indexing, data conversion, storage and other services.

a) A contract should be agreed with the service provider that details the services that are to be used.

b) If the contract does not require that the contractor comply with all relevant recommendations of this Technical Report, the user's inspection procedures on services provided should be such that no assumptions are made regarding the completeness, quality and accuracy of the services.

The procedures and recommendations in this subclause cover any type of service, including those provided on a facilities management basis, and are intended to ensure:

— that where work is carried out by a service provider, the procedures for the demonstration of authenticity of the resulting information will be the same as if the work had been done wholly within the client's organization;

— that the client can demonstrate compliance, many years after the event, even if the service provider has ceased to trade.

Where work is undertaken off-site, details of the procedures used in the transfer of information and/or media from the client to the service provider, and from the service provider to the client, should be documented.

If the service provider uses procedures which comply with the Policy Document, the client should hold a copy of, or have access to when required, the service provider's compliance documentation.

6.14.2 Procedural considerations

In ideal circumstances, where the service provider can demonstrate the implementation of procedures which comply with the Information Management Policy Document, the contract need only confirm this situation, and contain agreed procedures for checking compliance.

Where the service provider operates in compliance with agreed procedures, the contract should include a statement detailing the extent to which the procedures are implemented and audited.

The following defines procedures and processes that need to be reviewed and included within the contract as appropriate.

— The client should check that the service provider can produce output to agreed acceptable quality standards.

— The client should check that the service provider can process a sample of input material to produce output on the proposed media and in the proposed format and which can be successfully loaded on the client's target system. This sample should be retained.

— The client should check that the service provider can supply a copy of the audit trails of the processing undertaken in a readable form.

— Where indexing services are provided, the client should check with the service provider that the proposed indexing data accuracy requirements are acceptable and documented.

— The client should check that the proposed location of the work is acceptable and meets security criteria appropriate to the client's needs.

— The client should check that the proposed procedures and processes involve no greater risk of damage to the client's material than the client's procedures.

— The client should check that, where the material to be processed is unique or particularly valuable, effective fire detection and prevention systems are implemented at the proposed production location.

— The client should check that, where security of the material to be processed is important, the service provider should vouch for the trustworthiness of the intended operational staff. It is an advantage if all employees of the organization sign a confidentiality agreement as part of their conditions of employment.

— Where documents are sent for scanning, the service provider and client should make arrangements for documents to be accessible to the client whilst they are away from the client's premises.

6.14.3 Transportation of documents

Where documents are physically moved from the client's to the service provider's premises, opportunities exist for their loss or damage. Procedures need to be agreed to ensure that this risk is acceptable. Each shipment of material to/from the client and the service provider should be accompanied by a control document stating the identity and number of items included.

All material being shipped should be adequately packed to avoid risk of damage in transit.

The recipient should promptly check received material against the despatch document and advise the sender of discrepancies as soon as practically possible.

Transportation services may be provided by the user's own organization, by the third party, or by an independent courier.

Third parties providing transportation services should be organizations demonstrably meeting the quality and reliability criteria of the client.

Notes should be taken of the date and time at which the material was handed over to the transportation service and the date and time at which it was received by the service provider, and signed by the person handing over and receiving the material, respectively. The same process should be implemented on receipt of returned material.

6.14.4 Use of trusted remote archives

A secure means for detecting any tampering with a data file, or for verifying the contents of a file, is to store a copy of the file with a Trusted Remote Archive.

If such an approach is taken, an authenticated copy of the electronic file should be made and delivered either physically or electronically to the third party, using secure means.

The third party should follow the relevant procedures for the storage of information as recommended by this Technical Report, and should be able and prepared to demonstrate, in the same manner as the owner, the effectiveness and security of its services.

NOTE Security requirements for Trusted Remote Archives are frequently more stringent than those for the organization whose information they are storing.

Where digital signatures are used for authentication, instead of storing digital signatures in its own system, the organization may transmit the digital signature of a file to a Trusted Remote Archive. The third party will store the digital signature in secure conditions, such that it may be retrieved later.

6.15 Workflow

Some information management systems incorporate a workflow capability. Such systems provide the procedural automation of business processes, by the management of the sequence of work activities and the invocation of appropriate human and system resources associated with the activity step.

Where workflow systems are implemented, operational details (such as flow diagrams), process definition classifications and process definitions should be documented.

Process definition life cycles include:

— definition;

— development;

— implementation;

— withdrawal;

— modification.

All data (databases, audit trails, etc.) held on the workflow system should be reviewed for retention requirements and, where applicable, stored in compliance with the Information Management Policy Document.

Where changes to the workflow system are implemented, change control procedures should be implemented to ensure that stored information is not lost during the procedure.

Where *ad hoc* workflow is implemented (i.e. one in which the rules may be modified or created during the operation of the process), a full audit trail of the process should be kept together with the identification of personnel who performed the changes to the standard workflow procedures.

6.16 Date and time stamps

Procedures for the regular checking of system clocks for accuracy concerning date and time should be documented. Any errors should be corrected and any actions taken documented.

If the clocks are changed on a seasonal basis, e.g. summer time, then the procedures to be followed should be documented.

Only authorized personnel should be able to change system clocks.

Where there is a particular need to demonstrate the accuracy of date and time stamps, the use of trusted third-party service for this may be considered. Where trusted time is used, procedures for demonstrating the integrity and authenticity of a time stamp and its binding to a particular piece of information should be documented.

6.17 Version control

6.17.1 Information

In some applications, documents may be subject to change. Typical of such applications are those implemented for controlling technical drawings in drawing offices. Several different versions of a document may develop over a period of time, each document being allocated a version number. It is important in such applications to maintain each version as a separate document, and also maintain the link between the versions.

Where changes are allowed to stored image files, the procedures for authorizing and implementing such changes should be documented.

Documentation regarding any requirement to retain previous versions of such files should be available.

6.17.2 Documentation

A version control system may be implemented to ensure that the relevant version of any compliance document can be identified for any time in the life of stored information. A version control procedure should be established for all documentation.

Superseded versions should be kept for at least the same length of time as that for which relevant information is maintained.

Records of this maintenance are required so that the policies and procedures which were in force at the time of its capture and since that time can be described and attested to. If this is not done, there is a risk that the integrity of the information might be successfully compromised. For example, if it is not possible to be certain of the scanning procedures used to capture a document image several years old and the storage procedures

followed in the years since its capture, then it may be difficult or impossible to refute a challenge concerning the authenticity and integrity of the information.

6.17.3 Procedures and processes

All changes to procedures and/or processes should be implemented in accordance with an approved change-control procedure.

6.18 Maintenance of documentation

Compliance with the Information Management Policy Document requires the availability and use of specified documentation. Procedures for the maintenance of this documentation should be included in the Procedures Manual. Maintenance procedures should include the keeping of records of this maintenance.

Maintenance is required because, over time, requirements will evolve and technologies and legislation will change. In some cases, it will suffice for maintenance efforts to be driven by recognition of changes, on an *ad hoc* basis. Additionally, typically for more important documents, a routine regular review will be appropriate.

Procedures for ensuring that documentation is kept up to date should be documented.

These documents should be subject to records management disciplines which are at least as good as those applied to the organization's other vital business records.

In particular, whenever one of these documentation items is revised, a copy of that item prior to the change should be kept at least as long as the information to which it relates.

The storage of this documentation should allow for appropriate authorized parties (e.g. auditors) to identify and retrieve all the documents in force on any required date.

Documentation may be stored electronically in the Information Management System, subject to the same controls as included in this Technical Report, as paper or microform in secure locations, or as any combination of these.

The policy adopted for the storage of compliance documentation should be documented in the Policy Document.

In most cases, it will be desirable for changes to be documented in a way which allows an interested party to track the changes between versions. This can be implemented by recording a simple change history for each part of the documentation.

7 Enabling technologies

7.1 General

This clause deals with technology-related topics which are relevant to this Technical Report, including:

— System Description Manual (see 7.2);

— storage media and sub-system considerations (see 7.3);

— access levels (see 7.4);

— system integrity checks (see 7.5);

— image processing (see 7.6);

— compression techniques (see 7.7);

— forms overlays and forms removal (see 7.8);

— environmental considerations (see 7.9);

— migration (see 7.10);

— information deletion and/or expungement (see 7.11).

7.2 System Description Manual

A description of hardware, software and network elements that comprise the system and how they interact should be included in the System Description Manual.

Details of system configurations should be documented.

Details of all changes to the system should be documented. Such documentation should include details of any processes implemented to effect the change.

The System Description Manual should be structured so that details of the system at any time during the period of its use may be readily accessed. This may be achieved by creating a new version of the Manual every time there is a change, such that it is possible to access a clear description of the system as it was at a particular time in the past.

For systems already in operation, information stored on the system prior to the achievement of compliance with the Information Management Policy Document cannot be considered as meeting its provisions unless the controls and procedures described in this Policy Document have been in place from the time of storing that information.

The user should assess whether the elements of the system conform to the requirements of relevant National and/or International Standards. This enables system auditors to check the performance and reliability of the system against these Standards.

7.3 Storage media and sub-system considerations

The risk of stored image files being modified inadvertently or maliciously varies with the type of storage sub-system and medium. The ability to detect any such modifications also varies. For example, where write-once media is used, it is not normally possible to modify electronic files once stored, as any such modification would have the effect of destroying at least some data, resulting in files being corrupted, if not made totally irretrievable. Conversely, in the case of systems which use on-line storage, unauthorized modification, which is typically managed by access control, can never be totally guaranteed.

Image files stored on magnetic disk and other random access rewritable media may, in principle, be modified. With such media, the risk of modification is less to do with the medium itself than with the controls that are implemented by the storage sub-system and by the access software. The ability to alter files requires read-write access, and well-designed systems have controls to prevent unauthorized read-write access. Users with read-only access are unable to modify the files. This alone is unsatisfactory unless the system also maintains a secure record of all read-write accesses. In a system where there are very frequent file modifications, there may be a substantial overhead to record these modifications, but if a record is not kept, it might prove impossible to detect any unauthorized alterations by a skilled hacker or by anyone with the appropriate access privilege.

In the case of rewritable serial media, such as magnetic tape, unauthorized tampering can be more difficult than with random access media, since if the file which is modified is not the last file stored on the medium, then all following files need to be copied and rewritten. Once the medium is off-line, it could be tampered with more easily if an attacker were able to gain access to it. The issue of physical security of the off-line medium and access control while it is on-line is important.

The point in the application processes at which electronic files are requested by the software to write to storage should be documented.

In all cases, regardless of the actual medium, or the system environment in which the medium is used, procedures should be implemented to prevent modifications being made to stored information without detection. These procedures may involve use of electronic/digital signatures and/or copies that are stored in different locations; possibly involving trusted third parties.

In systems, which do not include facilities which, in the course of normal operations, would automatically detect unauthorized alteration to or removal of files, users should conduct random checks to verify that files which have been frozen have not been altered or removed.

Where write-once media is used, consideration should be given to the retention period of the information being stored. Where practical, information with differing retention periods should not be stored on the same physical piece of medium.

7.4 Access levels

Detail of all levels of access available on the system and procedures for their use should be documented. These levels are usually available as follows:

— System Manager;

— System Administrator;

— System Maintenance;

— Authors or originators;

— Information Storage and Indexing;

— Information Retrieval.

Only staff with the relevant access rights should be permitted to enter or amend stored information.

System access rights should be granted only after the member of staff has successfully proved his or her competence.

7.5 System integrity checks

7.5.1 General

Facilities should be provided within the system to ensure that the integrity of stored information is preserved throughout the system, including during its transfer to and from the storage media.

A suitable approach is to utilize a checksum calculated immediately after the information has been captured. This technique ensures that any errors in file transfer between sub-systems may be detected automatically and with certainty. Such a method on its own does not cover the possibility of malicious manipulation of the information between the time of capture and the time of committal to the storage media. Such manipulation could be accompanied by the calculation of a new checksum if the checksum algorithm were known. To deal with this eventuality, other procedures are required. A simple method is to write each checksum to the audit trail after calculation.

To protect stored information from malicious software, appropriate protection software should be installed and kept up to date.

Where appropriate, hardware to protect the system from power failure (e.g. an uninterruptable power supply) should be installed.

7.5.2 Digital and electronic signatures (including biometric signatures)

Digital and electronic signatures offer the possibility of demonstrating that retrieved information is exactly what was stored, and of confirming the identity of the people involved. The implementation of these signature systems usually requires the cooperation of both parties. Signatures are either created with signature digitizing devices (electronic) or using a key (digital), and are associated with the electronic file. The retriever then may use the signature to verify the identity of the original signatory and, with some signature systems, the integrity of the file. This applies to storage, workflow or transmission, whether real-time or store-and-forward transmission systems are used. Signatures should be used in applications where it is important to be able to confirm the integrity of a received file and the identity of the sender. Signatures should be stored securely. Access to signature files, keys and algorithms should be allowed only to authorized personnel.

Electronic signatures are usually stored within the files to which they are bound. Digital signatures need not necessarily be stored with the file to which they pertain, but it should always be possible to identify which file a particular digital signature is associated with, and vice versa.

Processes used for the creation of electronic signatures should be documented. These processes should include mechanisms for verifying the true identity of the person prior to that individual being enrolled as a document signatory.

Digital signatures are useful techniques for integrity and binding information to a specific entity (frequently an individual). However, care needs to be exercised over placing too much reliance on a digital signature without understanding the underlying technology and the risks associated with it. Different signatures can have different strengths, and rigorous security measures are needed in order to prevent a valid, but false, digital signature being generated. Also, where long-term storage of electronic files is envisaged, computer technology available in the future may be able to compromise digital signatures without detection.

Digital signatures should be used in applications where it is important to be able to confirm the integrity of a stored or transmitted file and, where appropriate, the identity of the party concerned.

Processes for the generation of keys should be documented.

Digital signatures, keys and algorithms should be stored securely, and access to them should be allowed only to authorized personnel.

If a query is raised about the authenticity of an electronic file, signatures may be used as evidence in demonstrating that any file stored or received by transmission contains the same information as the original file. Processes to be implemented where a query is raised about the authenticity of a file containing a digital signature should be documented.

7.6 Image processing

To provide optimum image output, or improve recognition rates for an automated data capture process, post-scanning processes may be performed. Where post-scanning processes are performed, the effect on the image of each of these processes should be individually documented.

The term *post-scanning processes* is used to describe various image enhancement techniques using hardware and/or software that can singularly or independently have an effect on the presentation of image output and the size of the stored file. They can be installed either on a scanner workstation or on a network server.

The more common techniques include:

— deskew;

— despeckle/background cleanup;

— black border removal;

— forms removal.

Image processing facilities ought to be used with care. For example, the despeckle process may remove decimal points, altering the value of numbers.

Any processing performed on the digitized image should not affect the integrity of the image as a true facsimile of the original. To check that any image processing does not affect the integrity of the scanned images, a sample set of documents should be scanned with the image processing active and prints of these images compared with the originals.

Where image processing techniques are used, consideration should be given to storing images of the sample set of documents with and without image processing.

The effect of processing performed on a grey scale image prior to conversion to a black-and-white image should be checked for acceptability.

Speckle removal should only be used with particular care, and its use should be documented. Speckle removal results in the elimination of single pixels or small groups of pixels from a digital image, resulting in a subjectively cleaner image, but it cannot be relied upon just to remove noise from the image. With some kinds of document there is a high risk that information may be removed, e.g. parts of already broken characters, punctuation marks, or parts of fine detail in drawings.

If speckle removal is used routinely on images, then without explicit information on the identity of images to which it has been applied, it may be assumed subsequently that all images have had speckle removal applied. This could affect the ability to demonstrate the authenticity of these images, if any doubt was raised about the completeness of the images.

The use of speckle removal may be documented in the operator log, or elsewhere in the audit trail, or by using additional data associated with the relevant image.

Where it is important that there should be no loss of information in the scanned image, other than that due to the scanning resolution, there should be no image processing subsequent to the initial creation of the image file.

7.7 Compression techniques

The use of file compression techniques should be in accordance with the Information Management Policy Document. Such techniques may be applied to electronic files by the system prior to or during storage, to reduce the size of the files. This also tends to improve system performance.

The type of compression used is usually application dependent, though some systems may have built-in compression that the user has no alternative but to use.

Compression can use various mathematical approaches, but all may be classified into two classes, namely lossy or lossless.

The compression techniques used, and their lossless or lossy attribute, should be documented. The documentation should be quantitative and include the algorithm used to compute the extent of loss.

This information may be stored as part of the file or its related data, or via a separate log.

NOTE For example, in the case of image files stored in TIFF (and some other) format, the compression method is automatically stored within the image file.

Lossy compression techniques should be used with care. By definition, lossy techniques mean that information is removed from the stored information during the compression process, so that the decompressed electronic file may not be the same as the original file. This may make the demonstration of authenticity of such files more difficult: for example, on an image file, parts of text or drawings may be removed, being replaced by artificially generated data. Thus there may be risk in using lossy compression on files containing primarily text (including handwriting) or line drawings.